

Slowdown will undermine reform

France's new research minister faces tough challenges in introducing much-needed changes into the research system. Her difficulties are compounded by impending budgetary constraints.

"Sacrifice research, and the entire future of our country is compromised." So proclaimed French presidential candidate Jacques Chirac during this spring's election campaign. If elected, he said, he would launch "a national mobilization plan for research and innovation", and boost research spending from 2.15% to 3% of GDP by 2010. The thousands of scientific jobs created under the previous government of Lionel Jospin "fell far short of needs".

But 2010 is a long way away, and reality is biting now. By mid-September, Jean-Pierre Raffarin, Chirac's prime minister, must somehow find billions of euros from a sluggish economy to pay for his boss's lavish election promises. As a result, research is set to be sacrificed in the September budget. The finance ministry is reported to have called for steep reductions in public research spending and cuts in research jobs. These have been forcefully resisted by Claudie Haigneré, astronaut and minister for research and new technologies (see page 712). Last weekend, Raffarin promised that the budget would not be cut, but maybe the best Haigneré can hope for is a flat or miserly budget.

She has done the calculations. To reach 3% by 2010, and taking the most optimistic scenario, whereby private-sector spending on research would increase three times faster than the public spend, the former would need to climb by 8.6% every year until the end of the decade, and public spending by 4.2% annually. The government, with the industry ministry in the driving seat, hopes to encourage such incredible growth in research spending, in particular in the life sciences, by subsidies and tax breaks. But by failing to boost the public sector, the government risks a dangerous imbalance.

Seven weeks into her job, Haigneré cannot be blamed for the poor budget prospects looming. Given the political context it is doubtful

that any science minister could have won a decent budget. The big question about Haigneré, a political neophyte, is whether in future she will be able to defend research in the ferocious world of politics.

Each new science minister in France faces the same old challenge: transforming a top-heavy, overly complex, slow-moving ship into a nimble craft where money and people can circulate more freely and be concentrated on the best science. In that context, Haigneré may be exactly the kind of minister France needs. Rather than rearranging deck-chairs on a sinking ship — almost all the key science indicators show France lagging behind most of its European neighbours — Haigneré seeks to pursue a productive strategy in which the ministry and the research community identify critical areas where tangible progress can realistically be achieved.

The areas emphasized so far are those needed to transform France's centralized and feudal system from within, by injecting flexibility: freeing up the vital force of young scientists by concentrating resources on giving independence to the brightest minds with the best projects, and seeking to redistribute more funds competitively on the basis of excellence, rather than sharing among existing labs. Haigneré wants to put an end to premature recruitment to full-time posts by creating an interim postdoctoral phase offering attractive conditions to young scientists who must then prove their worth.

Impending research spending will be a far cry from Chirac's pledge of a "historic commitment going beyond anything done in the past". This will not only compromise France's future, but also make Haigneré's already difficult task of shaking up France's rigid and inflexible science even harder, as there will be no sugar to help swallow the necessary pills. ■

Save starry nights

Cities should stop lighting up the heavens.

From the Ponte di Rialto in Venice on a clear night you can see every star in the constellation Ursa Minor. In nearly all of the world's other cities, from New York to Sydney, you can't. Venice has maintained the privilege of a starry sky because it has no traffic, and because public lighting has, until now, always been subdued, in keeping with the city's romantic architecture. But even Venice's night sky is now threatened by a new mayor who supports the introduction of mercury street lighting.

Fortunately, hopes that light pollution will not worsen are not completely forlorn. Over the past decade, the small but passionate lobby for legal standards in outdoor lighting, spearheaded by the International Dark Sky Association in Tucson, Arizona, has witnessed the passing of lighting laws and ordinances in several US states, counties and cities, and more recently in Europe, notably in Lombardy and other Italian regions. The Czech government passed a lighting control law in February this year; it is the only such national law in existence.

The stars are our rightful heritage. Moreover, light pollution interferes with ground-based astronomy, disturbs human sleep and upsets nocturnal wildlife. None of that moves the hearts of legislators,

however: they tend rather to be swayed by the economic argument that light thrown sideways and upwards is wasteful of energy.

Most of the laws relate only to state-funded lighting, and set a limit on light pollution, rather than banning it completely, even though there is now no shortage of companies that can provide pollution-proof fittings. More and tougher laws are required to meet even the modest goal of not worsening light pollution levels, and this requires a lobby that will be taken seriously. The International Dark Sky Association has turned to UNESCO (the United Nations Educational, Scientific and Cultural Organization) for support. UNESCO has given its powerful 'world heritage' status to 730 cultural and natural sites worldwide — Venice among them — and initially wanted to get the night sky on the list. But its own rules require that the heritage to be protected belong to a particular government.

UNESCO is apparently considering a proposal to create a list of heritage sites that nobody owns, such as the oceans and sky. Human health might not be damaged by a lack of starlight, but the quality of urban citizens' lives will be significantly diminished if the night sky is available to them only in planetariums. ■





On a mission
Former astronaut
gets to grips with
French research
p712



Data drought
Storm brews over
India's weather
forecasting
p713



Bitter pill
Britain's medical
research faces
funding crisis
p714



Gloomy outlook
Brown haze sparks
fears of climate
change
p716

Senators attack Pentagon over weak state of defence research

Geoff Brumfiel, Washington

Prominent US senators have charged the Department of Defense with doing too little to halt the brain drain from laboratories that carry out research for the US military.

In a letter to Donald Rumsfeld, the defence secretary, seven members of the powerful Senate Committee on Armed Services, led by



Joe Lieberman:
blasts Pentagon.

Joe Lieberman (Democrat, Connecticut) and James Inhofe (Republican, Oklahoma), expressed frustration with what they see as the Pentagon's tardy efforts to reinvigorate the laboratories, which have been losing talent and prestige for years.

In particular, the senators attacked the Pentagon for failing to exploit special hiring systems, created by Congress, to attract young scientists back to the labs. "We have become increasingly dissatisfied with the systematic underutilization of these provisions," the 1 August letter complained.

The Pentagon's research system includes the three major service laboratories — for the Army, Air Force and Navy — as well as a number of specialized testing centres, employing

a total of nearly 15,000 people. The Naval Research Laboratory based in Washington DC, in particular, was once a renowned centre of international excellence in several disciplines. But since the end of the cold war, the laboratory's workforce has shrunk by about a third, and many of the staff are now nearing retirement, says Vice Admiral Paul Gaffney, who was chief of naval research from 1996 to 2000.



Dissatisfied:
James Inhofe.

Civil-service hiring rules have hampered efforts to attract new staff, says Lance Davis, an official at the National Academy of Engineering who formerly oversaw the labs from the Pentagon. Lab directors have to post openings for long periods of time and Pentagon officials must sign off each hire — delaying job offers for months, Davis says. "Directors who run billion-dollar laboratories have no control over who they work with," he adds.

To try to fix matters, Congress has repeatedly legislated to create pilot programmes that allow the Pentagon to streamline hiring and give the labs more salary flexibility. The senators also asked for laboratory directors



All at sea? The Navy's research laboratory has faced problems attracting young staff.

to be granted authority to hire and fire without the approval of Pentagon bureaucrats. "Congress gave the Department of Defense this ability and they haven't done anything with it," says one Senate staff member.

Cynthia Colin, a Pentagon spokeswoman, said she could not comment as officials had not had time to review the letter and formulate a response. ■

Online cheats leave Asian students facing paper exam

Erika Check, Washington

Tens of thousands of students from four Asian countries are to be barred from taking part in the Graduate Record Examination (GRE) general test online after their compatriots were found cheating in it.

Administered by the Educational Testing Service (ETS), the exam tests skills, such as mathematics and English-language ability, of graduate school entrants in the United States. The ETS says that students in China, Hong Kong, Taiwan and Korea will have to take paper versions of the test this autumn, after officials uncovered large-scale cheating on the verbal section of the online GRE.

The ETS says it unearthed two websites

where students had posted questions and answers that they had memorized while taking the test. These questions had been used many times in the online test. Participation in the test is required by most US graduate schools as part of candidates' admission applications.

The ETS will not say how many students accessed the websites. But its spokesman Tom Ewing says that the cheating was confined to the four countries where students will now have to take the paper test. More than 55,000 students took the computer-based GRE last year in the four countries, and Ewing says that a "substantial number of students" must have cheated, as average scores for the verbal

section rose by as much as 100 points during a single testing cycle. Each section is graded on a 200–800-point scale.

But some graduate students are concerned that the paper GRE will not be directly comparable to the online version to be taken by students in the rest of the world. "The ETS should investigate and clarify the extent of cheating, so that students would not be suspected unnecessarily," says Ying Wong, a graduate student in psychology and member of the graduate student council at Stanford University in California. "It would be very damaging to the general Asian student population's morale if they feel their reputation cannot be cleared," Wong says. ■

GETTY IMAGES

AP

T. GIPSTEIN/CORBIS

French minister plans shake-up for research

Declan Butler and Sally Goodman, Paris

Claudie Haigneré, who became minister for research in France's new centre-right government in June, was the nation's first female astronaut and is the only woman qualified to pilot the International Space Station's Soyuz lifeboat back to Earth.

Now she's on an emergency mission of a different sort. She is aiming to rescue the country's research system from a set of problems that many observers say are preventing France from keeping its best young talent or from punching its weight internationally.

Haigneré's aim is to rebuild relations between her ministry and the research community, and then bring about change with the backing of scientists. "I'm aware of the imperfections of the system, but we can't change it with a magic wand," she says. She has pledged not to micromanage the various agencies supervised by the ministry, but rather to render them accountable by evaluating their performance in reaching set goals.

Haigneré is a calm, unassuming character. But her professional track record suggests that this political debutante may just have the nerve and tenacity to get the job done.

She topped her class at the University of Dijon medical school, and practised in rheumatology and sports injuries before gaining a PhD in neuroscience and a research position at a lab run by the CNRS, France's national research agency, in Paris. In 1985, Haigneré was chosen from 1,000 applicants to become an astronaut. While training, she led medical research at the French space agency.



New mission: former astronaut Claudie Haigneré is now at the helm of France's research ministry.

Haigneré has pledged that her approach will be forged by the "patience, determination, continuity of effort, and teamwork" that she learned from preparation for space missions. She is seen as a reformer — but one in a very different mould from the geophysicist Claude Allègre, who attempted a radical overhaul of French research, without the backing of researchers, during his tenure as minister from 1997 to 2000 (see *Nature* **404**, 421; 2000).

One of Haigneré's goals will be greater autonomy for young researchers. Almost 75% of full-time researchers in France work in the laboratory where they obtained their PhD. Because most funds are spread thinly across

laboratories, rather than going to individual investigators as competitive grants, young researchers often find themselves subservient to powerful senior scientists. As a result, critics say, many of France's best young talents disappear abroad and do not return.

Haigneré says she would like to break this pattern and offer the best young scientists higher salaries, larger grants and better working conditions. She favours the creation, for the first time, of a full-blown postdoctoral system in France, but does not plan to challenge the civil-servant status and good job security currently enjoyed by public researchers.

The time may be ripe for such a pragmatic approach to reform, says Daniel Louvard, director of the Curie Institute in Paris. What's needed, he says, is not wholesale change, but rather a set of discrete measures that will "encourage young scientists to take risks and give them good reasons to work in France".

But Haigneré's ambition may be hampered by the likelihood that next year's science budget will be unchanged or even reduced, potentially jeopardizing an agreement between scientists and the previous government to increase the number of research positions.

Researchers are already complaining that President Jacques Chirac has reneged on a promise, made before his re-election in May, to increase total spending on research and development from 2.15% of France's economy to 3% by 2010. Haigneré says most of this increase will come from industrial spending, but admits it will be difficult to meet the target unless the public research budget also grows. ■

Scientists seek safety in secrets of the soundbite

David Adam, London

A scare over the safety of a widely prescribed medicine is gripping the nation. As a prominent scientist, you're being interviewed live on television. Then comes the dreaded question: "Is it safe?". How do you respond? Insist that it is and risk sounding dismissive? Reach for a reassuring statistic and be accused of hiding behind numbers? Or admit that "we just don't know", and satisfy nobody?

Help may soon be at hand. Late last month, a group of scientists, journalists, politicians and risk-management experts met at London's Science Media Centre to discuss how to handle this question. The centre, launched in April by the Royal Institution to improve representation of science in the media, intends to use the results in a leaflet on how best to communicate risk. "When scientists are asked the question 'is it safe' and realistically it is, then we want to give them some options to draw on,"

says Fiona Fox, the centre's head.

Three volunteers agreed to undergo mock interviews at the hands of BBC science correspondent Pallab Ghosh, who confronted them with a barrage of questions on issues that have captured the media's attention in the past six months: the risk associated with the combined measles, mumps and rubella (MMR) vaccine, hormone-replacement therapy (HRT) and railway safety.

Opinions varied as to how the risks involved should be communicated. Some said that a humble "in my opinion" or an "I am confident" help to suggest relative safety. Others recommended relating the risk to other activities — "I've already done five riskier things today," said Colin Berry, a pathologist at the Royal London Hospital. And Evan Harris, the Liberal Democrat science spokesman and Member of Parliament for Oxford West, said phrases such as "if you think sex is risky you should

try lifelong celibacy" can help to make clear the benefits of taking a particular risk.

Other phrases that went down well with the audience included "breast cancer is a risk of being alive", and "I can understand that people are worried but...". For MMR, the participants liked the idea of stressing the fact that several studies had looked for a link with autism, but none had found one. But scientific jargon, such as "uncertain in a compound way", is best avoided, the audience said. One participant was advised against describing the risk associated with rail travel as "unquantifiable" — even if the statement is technically correct.

As well as preparing suitable answers, Ghosh said that scientists should bear in mind that they are talking to the audience, not the interviewer. But Harris added that addressing all of the audience is impossible. "There are some people that you can never and will never persuade, whatever the

Climate model under fire as rains fail India

K. S. Jayaraman, New Delhi

As much of southern Asia faces its worst drought for 30 years, the Indian government is being heavily criticized for its reliance on a statistical model to forecast the monsoon.

The India Meteorological Department (IMD), based in New Delhi, uses the model to issue a forecast of the monsoon every spring. This May it issued a prediction saying that rainfall would be 101% of normal. But by 6 August, this forecast remained way off the mark, with about three-quarters of the country having received very little rain.

Most of India's rain falls during the summer monsoon season between June and September, with July typically accounting for half of the annual total. But this season has been unusually dry, and last week's heavy rain and flooding in some areas are unlikely to save the harvest, most observers say. An accurate forecast of the impending drought would have allowed farmers to switch to more resilient crop varieties, as well as giving the authorities more time to prepare for a potentially poor harvest.

The IMD's statistical model is based on past weather patterns. It tracks 16 factors thought to have some influence on the monsoon, ranging from the amount of snow lying on the Himalayas to the status of the El Niño climate pattern in the eastern Pacific Ocean.

Many researchers point out that the prediction of extreme events — such as the failure of the rains — will always be challenging for this type of model. But some are now



Poor prediction: the Indian government failed to forecast this year's drought, the worst for 30 years.

questioning the IMD's choice of criteria, and others say that the government should switch to more sophisticated general-circulation models (GCMs). These calculate the evolution of the ocean-atmosphere system by combining dynamic equations with data on initial conditions.

"The real challenge in long-range monsoon forecasting is to predict the extreme drought events," says J. Srinivasan, a climatologist at the Indian Institute of Science in Bangalore. "Because we had normal or above normal rainfall for the past 14 years, we may have placed more faith in statistical

models than we ought to have done."

But the government is playing down the failure of this year's prediction. "It is unfair to blame the model on the basis of one failed forecast in 15 years," says Valangiman Ramamurthy, secretary of the Department of Science and Technology (DST), which oversees the IMD. "We will review the model at the end of the season." He concedes that an effort is already under way to refine the model, however, noting that four of its 16 criteria were changed two years ago.

S. M. Kulshrestha, former director of the IMD, notes that a high degree of variability is intrinsic to the monsoon. "You can throw the model away, but there is nothing to replace it with," he says.

Most researchers see GCMs as the best long-term bet for monsoon prediction, although GCMs cannot yet predict monsoon failure accurately. But they believe that these predictions will improve as more observations are made, the models become more detailed and computers increase in power.

GCMs are already showing some promise. This year, for example, the European Centre for Medium-Range Weather Forecasts model predicted normal Indian rainfall based on conditions on 1 May, but below normal based on 1 June conditions. "This indicates that this year's below normal rainfall occurred on account of changes in atmospheric circulation during May," says Srinivasan.

Officials still doubt whether GCMs can replace the statistical model in the short term. There are lots of GCMs "but no consensus about which one to use," says Dev Raj Sikka, head of the DST's climate-research programme. Nevertheless, Indian scientists agree that this year's unusual monsoon conditions will provide useful data on which to base future predictions.

evidence, so I try and give answers to the large number of people who are swayed," he said.

Statistics provoked the most heated debate. Some argued that figures bring much-needed perspective; others said that

they cloud the issue. For example, media coverage of the suspension of a recent US HRT trial focused on the apparently alarming 26% increase in the incidence of breast cancer, although this is equivalent to just 8 extra cases per 10,000 women.

Some scientists have already proved to be capable of putting such figures into context. In a 10 July interview on BBC Radio's *Today* programme, David Purdie, an obstetrician and gynaecologist at the Hull Royal Infirmary, discussed with interviewer John Humphrys the apparently huge increases in risk posed by HRT trials. The figures, he said, are an example of "the Judas factor" — a small number of cases can cause an alarming-sounding percentage if the total number of samples is small. "Christ was betrayed by 8.5% of his disciples," Purdie told Humphrys. "But when you consider how many actually did the job, it was just one out of 12."

♦ www.sciencemediacentre.org



Top projects suffer as medical funding falters

Natasha McDowell, London

British biologists funded by the Medical Research Council (MRC) are facing shortfalls in cash, after investment in a series of new projects has left the council unable to award grants to many top-ranked projects, *Nature* has discovered.

The MRC, which distributes British government funds for medical research, generally provides money for all projects given the top alpha-A rating by the council's peer-review board. But at a meeting last month, the council decided not to fund 36 alpha-A projects, and to delay funding for a further three. Forty-five alpha-A projects received funding. "It has been a difficult year for applicants and we are sorry not to have been able to fund more," says George Radda, the MRC's chief executive. It is likely that funding will be tight next year, he adds.

An MRC spokesperson says that the shortage is due to the council's investment in large new projects such as the UK Biobank, a database of the genetic details of half a million British people. The £45-million (US\$69-million) project, which was launched in April, is funded by the MRC together with the Wellcome Trust and the Department of Health. The data will be stored alongside medical records and lifestyle information, and will be made available to both academic and commercial groups.

Such projects have the backing of many

British researchers, but those who have been denied funding are angry at the way the MRC has managed the situation.

"It's extraordinary that they have let it happen," says one researcher, whose five-year programme grant, which covers salaries for postdocs, research assistants and lab supplies, was not renewed despite being awarded the top peer-review ranking. "They should have made sure they had enough money to fund all the alpha-As. The message it sends out abroad, to potential postdocs, is that UK funding is fickle, irrespective of the quality of the science."

Researchers who have not been granted funds are also disgruntled at being left in the dark about why they missed out. The council ranks alpha-A projects and uses the results to decide which to fund, although these rankings are not released. A decision to delay decisions on a previous batch of alpha-A applications, made at an MRC meeting last October, has angered applicants further. Some researchers were reassured at the time at receiving alpha-As, only to be denied funding nine months later.

"The message it sends out to young researchers is that there is no job security and no future in research," says another MRC-funded researcher.

Many researchers contacted by *Nature* say they are worried that university research groups could be particularly vulnerable. The MRC has other funding commitments, such



MRC head George Radda says the council regrets having been unable to fund more projects.

as salary costs, at the various research centres that it runs. The MRC says it cannot provide details on which projects have been funded, but stresses that all decisions were based on scientific merit.

♦ www.mrc.ac.uk

Lessons in research aim to win pupils over to science

Quirin Schiermeier, Bremerhaven

For 22 schoolchildren in the north German seaside town of Bremerhaven, the new school year has begun in unusual fashion. Every Tuesday and Thursday, they swap their classrooms for lecture theatres and labs at the Alfred Wegener Institute for Polar and Marine Research (AWI) on the coast of the Wadden Sea.

There they are taught biology, chemistry, physics, mathematics and English. The lessons bear little resemblance to the way that German 16–18-year-olds are usually taught science. No school bells interrupt the day-long projects; theoretical classes and practical sessions are closely intertwined; and there are no fixed timetables. Guided by science teachers and AWI researchers, the pupils will, over the next three years, learn the principles of scientific experimentation and how to use research equipment by carrying out field research and lab work on their own.

The scheme is the most ambitious of a recent series of attempts to stem the decline in the number of students choosing science at school and university. Some German universities and research institutes are already offering taster courses and internships for schoolchildren, but the AWI's attempt to integrate science education into the day-to-day work of a research institute is the first of its kind in the country.

The students will learn basic techniques in Earth sciences, such as how to test for the presence of various substances in water samples. They will then apply these to their first project — a study of the Wadden Sea ecosystem. The area is Europe's largest coastal wetland, flooded at high tide and dry at low tide.

AWI researchers have promised to do their best to support the scheme. "Once the students master the basic mathematical and scientific tools of the trade, there are many ways to let them participate in research

work," says AWI oceanographer Ulrich Bathmann. Bathmann has compiled a list of possible teaching topics, including carbon cycles, species diversity in phytoplankton and atmospheric processes.

Science education experts are enthusiastic about the idea of bringing schools and research centres closer together. "This is an exciting new approach," says Dagmar Schipanski, science minister of Thuringia and president of the German conference of education ministers. "Since reunification, I have observed the decline of young people's interest in science with great concern."

A few similar schemes have been run in Denmark and the United States. And Walter Kröll, president of the Helmholtz Association of National Research Centres, to which the AWI belongs, plans to extend the project. "Teaching laboratories similar to the AWI's will be created at all Helmholtz centres," he says.

♦ www.awi-bremerhaven.de/index-e.html

Air to stay? Brown haze settles on many parts of Asia, including densely populated cities such as Kuala Lumpur.



Brown haze darkens outlook for climate across Asia

Nairobi As India experiences one of its driest monsoon seasons for decades (see page 713), a study has suggested that atmospheric pollution is likely to make such droughts more common in some parts of Asia.

The report by the United Nations Environment Programme focuses on aerosols — small smoke particles from cars, industry and vegetation fires that create a thick brown layer of haze over some of Asia's most densely populated regions. The haze reflects sunlight, cooling and drying the areas it covers, which could decrease monsoon rain in some areas of central Asia by up to 40%, the authors say, while increasing rainfall in the southeast of the continent. The report also shows that the haze spreads further than was originally thought, reaching beyond populated areas.

The study is based on data from the Indian Ocean Experiment, an international effort involving scientists from Europe, India and the United States. But researchers say that they are a long way from understanding the effect of aerosols on global climate. Little is known about how aerosols affect climate variables such as cloud formation and global wind patterns, making it hard to factor them into climate models. These issues will be discussed next month at an international conference on aerosols in Taipei, Taiwan.

♦ www.unep.org

French fusion reactor maintains record hot streak

Paris A fusion reactor in France has kept hot plasma stable for three-and-a-half minutes, almost twice as long as the previous record set in 1996 at the same reactor. Prolonging the stability of such plasmas is a key element in developing a commercial fusion reactor.

The doughnut-shaped Tore Supra reactor at the French Atomic Energy Commission's research centre in Cadarache, near Aix en

Provence, uses an electrical current, rather than fusion, to heat the plasma. The record was achieved by funnelling away excess energy through carbon tiles on the machine's floor.

Although the plasma was too cool to produce fusion energy, the results should help researchers to design larger fusion reactors that can hold plasma for longer periods. The know-how gained at Tore Supra will also feed into plans for ITER, the proposed international project to build an experimental magnetic fusion reactor.

Couple accused of lab theft return to work

San Diego A married couple accused of stealing scientific secrets from Harvard University have returned to their posts at Californian research institutes.

Jiangyu 'Jonathan' Zhu, a Chinese citizen, and his Japanese wife, Kayoko Kimbara, were arrested in San Diego in June. Federal officials allege that the pair took reagents and genetic clones from a Harvard molecular-biology lab where they worked in 1999, and sent them to a Japanese drug firm. The couple are currently on bail.

Internal reviews have now been conducted at the University of California, San Diego, where Zhu holds a postdoctoral position, and at the Scripps Research Institute in La Jolla, where Kimbara has a state-funded fellowship. Officials say that the investigations found no evidence of impropriety by the couple. Both are now back at work, although university officials say that Zhu will not be allowed to take part in laboratory studies.

Transgenic cotton gets mothballed after protests

New Delhi India's decision to embrace some transgenic crops suffered a blow last week when the government in the southern state of Karnataka banned the sale of genetically modified seeds. Protests by farmers in the area had culminated in the burning of fields of transgenic Bt cotton in the Davangere district, prompting the government to issue its ban on 10 August.

The Bt cotton contains a gene for a toxin from the bacterium *Bacillus thuringiensis*, which allows it to resist attack from its main pest, the bollworm. It is the first transgenic crop to be given a commercial licence in India. Permission to plant it was granted by the national government in March (see *Nature* 416, 468; 2002).

Activists say that transgenic crops will damage the genetic diversity of India's crops. The protesters, which include the Karnataka State Farmers' Association, also want a transgenics research unit run by biotechnology firm Monsanto removed from the Indian Institute of Science campus in Bangalore.

Lawsuit sounds the alarm over Navy sonar

Washington A powerful new underwater sonar system proposed by the US Navy would endanger whales and other animals and should be made illegal, according to a lawsuit filed last week by a coalition of environmental groups.

The National Marine Fisheries Service approved the use of the system in July, despite last December's admission by the Navy that use of a similar sonar system resulted in the death of at least six whales when it was tested in 2000 (see *Nature* 415, 106; 2002). The Navy says the new system operates at lower frequencies than the one that caused the deaths, and that sonar is needed to detect new ultra-quiet submarines that are currently under development.

The New York-based Natural Resources Defense Council, which leads the environmental coalition, insists that such a system will physically maim whales, dolphins and other sea life, and disturb their feeding, mating and migratory patterns.

♦ www.nrdc.org

Bones hint at dinosaur king's princely relative

Chicago Palaeontologists have debated whether *Tyrannosaurus rex* had a smaller cousin for over half a century. Now a find by a team from a small US museum may help to settle the issue.

The dispute began with the discovery of a tyrannosaur-like skull in Montana in the 1940s. Some palaeontologists argue that certain distinctive features of the skull, such as its finely serrated teeth, distinguish it as a unique genus, dubbed *Nanotyrannus*. Others say the skull is actually from a juvenile *T. rex*.

The new find, unearthed in Montana last month by a team from the Burpee Museum of Natural History in Rockford, Illinois, and named Jane after a museum sponsor, could resolve the question. The discoverers believe that the fossil is not a young *T. rex* because it has fused vertebrae, which are only seen in adult dinosaurs. But they say they need to free the skeleton from the ground to be sure.



Digging the king? Palaeontologists are split on whether this is *Tyrannosaurus rex* or a relative.

Troubled waters

The oceans around the United States suffer from overfishing and pollution, but current government regulatory structures only hamper attempts to fix these problems. Can two high-level commissions put things right? Mark Schrope investigates.

The complex bureaucracy behind the ocean policies of the United States is positively kafkaesque. The country's most important marine organization — the National Oceanic and Atmospheric Administration (NOAA) — takes much of the responsibility, but countless areas fall outside its remit. Most water-quality issues are handled by the Environmental Protection Agency (EPA), whereas the Department of the Interior handles undersea energy resources. Overall, US oceans are controlled by a morass of nine government agencies, the budgets of which are overseen by 44 congressional committees and subcommittees.

The oceans desperately need a more coordinated approach. Overfishing has sent many fisheries into decline. The populations of some species, such as certain types of west-coast salmon, have already plummeted. Pollution is also taking its toll — some researchers believe that it may be behind the decline of Florida's barrier-reef system, the third largest in the world, parts of which have lost 95% of their coral. But when researchers, environmentalists and state authorities try to address these problems, they often find their attempts mired in an administrative bog.

Fortunately, this approach is set for a shake-up. A government-appointed panel of experts is immersed in an exhaustive review of ocean policies, and a similar, independent group is conducting its own analysis. As their consultations come to an end, the two groups have begun the daunting task of assimilating what they have learned. Their recommendations, when published, will have repercussions from Capitol Hill to the edge of the continental shelf — and could provide a boost for



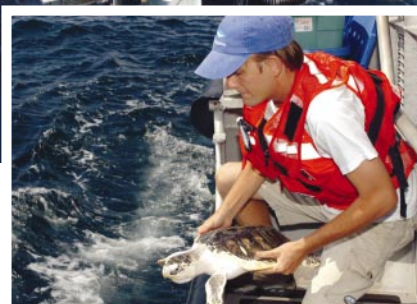
In the United States, energy resources (above), ocean research (right) and fisheries (below right) are all governed by different bodies.

ocean research. For those concerned about the oceans, the reports can't come soon enough. "If we don't do something about this crisis in the oceans soon," says Roger Rufe, president of The Ocean Conservancy, a Washington-based environmental group, "it's going to be too late."

Water birth

Reform is certainly long overdue. The last major review on US oceans policy — the Stratton Commission — was delivered in 1969. Named after the panel's chair, Julius Stratton, then chairman of the Ford Foundation charity, the commission called for the creation of a 'wet NASA'. NOAA was born, but the organization has never quite achieved what Stratton intended. President Richard Nixon declined to make the new body independent, partly because the report had been commissioned by his predecessor, Lyndon Johnson. And since its creation, NOAA has grown well beyond what Stratton envisaged. Its \$3-billion budget is now spread between everything from the national weather service to marine fisheries.

But NOAA does not take complete control of US oceans. A complex range of organizations set rules at state and federal level, giving rise to a situation in which fishing vessels, for example, encounter different regulations when they cross from coastal waters — which are under state control — into federal ocean territory. Matters were further complicated in 1983 when US waters were enlarged, mainly to gain control of more off-shore energy



resources. US territory was extended from the edge of the continental shelf, which in some place is less than 10 km from the coast, to 370 km offshore, but no new comprehensive legislation was passed to govern how the oceans should be managed.

Concrete plans to rectify this emerged in 2000, when two reform efforts began life. That May, the Pew Charitable Trusts, a large granting organization based in Philadelphia, announced the creation of the Pew Ocean Commission — a group of 18 experts, including environmentalists, researchers and government officials, led by Leon Panetta, who



Talking shop: commission heads Leon Panetta (far left) and James Watkins (above, right) hear from users of the ocean's resources.

had been White House chief of staff earlier in the Bill Clinton administration. Three months later, Clinton signed the Oceans Act into law. The act mandated a review of oceans policy, and in July 2001 the US Commission on Ocean Policy, a 16-member panel headed by James Watkins, secretary of energy under former president George Bush, was born. The Pew commission expects to deliver its recommendations early next year, and Watkins' commission aims for next summer.

Scope for change

Both groups are charged with creating a blueprint for a new ocean policy, and will have to examine everything from the structure of the agencies that govern the oceans to the problems of coral-reef decline and pollution. But the Watkins Commission's scope is broader, including analyses of commercial activities such as oil-drilling and shipping.

At the top of each commission's to-do list is a reorganization scheme for the agencies that oversee ocean issues, and the main question on the mind of observers is how radical the recommendations will be. "Are they going to suggest minor changes?" asks Michael Orbach, director of the Duke University Marine Lab in Beaufort, North Carolina, who has worked with both commissions. "Or are they going to be really revolutionary and suggest new and different ways to do things?"

The creation of a cabinet-level Oceans Department would be one possible revolution. Ocean-related segments of other bodies, such as NOAA and the EPA, would be encompassed by the new department. Its head — the oceans secretary — would have access to the president and direct involvement in high-level budget planning. A more minor reorganization would involve the reform of NOAA so that it focuses solely on ocean issues. Either course would provide a more focused approach to ocean management — but persuading administrators to let go of responsibility may be difficult. Any significant redistribution of duties would "be a very, very, very — and maybe that's not enough verys — difficult thing to make happen", says Bruce Molnia, the House of Representatives staffer who coordinates the Oceans Caucus, a bipartisan group

of representatives that monitors ocean issues.

A milder reform would be the creation of bodies to oversee ocean issues across different agencies. Such a body could be effective if it was given substantial "teeth", such as control over funds and the power to dole out responsibilities to the various agencies, says Ken Brink, a physical oceanographer at the Woods Hole Oceanographic Institution in Massachusetts and member of the Watkins Commission's science advisory panel. One such body — the interagency National Ocean Research Leadership Council — already does this for a subset of ocean issues.

Reaching a consensus is clearly going to be difficult. "It's not apparent to me that there is one predestined solution that everyone will agree to," says NOAA administrator Conrad Lautenbacher. "If it were that easy it would have been done a long time ago." But although the answer is far from clear, both commissions are at least considering signifi-

cant reforms. "The cry for new structures and management techniques has been coming in from everybody," says Watkins.

Pollution problems

As well as grappling with the fate of ocean policy, the two commissions are considering how to integrate land issues, such as reducing farm run-off, into the ocean-management process. "Land-based, non-point-source pollution is probably the biggest threat we have to coastal and nearshore environments," says Watkins. He cites the example of the Gulf of Mexico, which receives run-off laden with nutrients from fertilizers from 31 states and 2 Canadian provinces through the Mississippi River. This supports the growth of algae, the decomposition of which reduces oxygen levels to almost zero in bottom waters as far as 85 km from the Mississippi Delta. Marine life is killed or forced out of a huge 'dead zone', which can grow to the size of Massachusetts. But despite the scale of the problem, there is no system that allows effective cooperation between researchers, state and federal environmental bodies and the various agricultural agencies involved.

The concerned parties could look north-east to Chesapeake Bay for a solution. In 1983, federal, state and environmental bodies got together with the aim of protecting and restoring the plant and animal life of the bay, which straddles Maryland and Virginia. It was soon clear that changes in the bay itself were not enough, so the scheme — called the Chesapeake Bay Program and based in Annapolis, Maryland — was extended in 1992 to include states with tributaries that feed the bay.

Researchers from several local institutions and government agencies have played an active role in the programme. For example, models of water flow into the bay, the creators



Tried and tested: modelling of water flow into Chesapeake Bay (left) is helping local authorities to protect the area's ecosystem. Can a similar analysis be applied to the Mississippi Delta (right)?

of which say are the most advanced simulations of any estuary, are being used to set the water-quality standards that each state needs to achieve if the bay is to be removed from the EPA's list of impaired waters (L. C. Linker, G. W. Shenk, R. L. Dennis & J. S. Sweeney *Water Quality and Ecosystem Modeling* 1, 91–122; 2000). Lewis Linker, the programme's modelling coordinator and an EPA employee, says that the programme's final targets should be agreed upon by early next year.

Fishy story

Fisheries are another hot topic for the commissions. There are, for example, big holes in the data on fish stocks. The federal government lacks good information on the population size and health of around 68% of the stocks that it manages, and for those stocks whose status is known, 33% are overfished, according to a report published this year by the National Marine Fisheries Service (NMFS), a division of NOAA.

Fisheries management also gives cause for concern. Catch quotas are set by the NMFS, together with regional fishing councils. Stocks are managed one species at a time, often with little regard to other species, particularly those that are not commercially exploited. "You'll never have ecosystems management come to reality one species at a time because everything you do to one affects another," says Pew Commission member Pat White, a former executive director of the Maine Lobstermen's Association, based in York, Maine.

Any commitment to gaining more data and moving towards an ecosystem-level approach to fisheries management will mean more work for scientists. In Chesapeake Bay, for example, a government ruling on the use of shrimp nets outside the bay, created with the aim of reducing the number of other species that are caught in the nets, led indirectly to a reduction in the bay's population of blue crabs. Researchers think that one cause may be levels of croaker fish, a natural predator of the crab and a victim of shrimp nets, which increased after the new regulations were introduced. Croaker populations inside the bay may also have risen, triggering a fall in crab numbers. Effects such as these have prompted



All aboard: oceanographers hope to be given access to an enhanced fleet of research vessels.

NOAA to fund the creation of a detailed model of the bay's ecosystem, and better management of fisheries is likely to require similar models for other areas. "I think what we're developing is going to demand an even larger role for scientists than they have today," says Panetta.

Mobilizing the fleet

The commissions also have the power to give considerable boosts to other areas of ocean research, and dozens of scientists have testified at the commission's regional meetings. Ocean researchers will, for example, be hoping that the Watkins Commission backs infrastructure funding. Many scientists say that new research vessels are required, and parts of the existing academic fleet will need overhauling. A schedule to do this has been put together by the organization that manages the fleet — the University–National Oceanographic Laboratory System, based in Moss Landing, California — but it is unclear where the money will come from.

Others would like to see the commission lend support to endeavours such as the NEPTUNE project, a \$250-million plan to wire a tectonic plate off the west coast of the United States with an array of sensors and equipment to monitor its geology, biology and chemistry. And most researchers are lobbying for the Integrated Ocean Observing System, a network of buoys, satellites and ships that would provide a comprehensive source of ocean data (see *Nature* 416, 253; 2002), and which Watkins describes as one of the projects he expects his commission to support.

Advocates of ocean-exploration projects, which aim to study areas of interest rather than pursuing well-formed hypotheses (see *Nature* 412, 672–673; 2001), are also hoping for support from Watkins' commission. Last autumn, NOAA's Office of Ocean Exploration, which was formed in 2001, launched one of its first missions — a three-leg voyage that included studies of deep-sea corals and methane deposits. Advocates of ocean exploration have a long list of other projects, including missions to produce maps of the ocean floor detailing marine life, mineral resources and geological hazards such as

earthquakes. NOAA only has \$14 million to fund exploratory projects — well short of the \$75-million multi-agency projects that advocates have called for — but the approval of one or both commissions might help to persuade politicians to rectify this.

If any of these goals is to be achieved, the commissions will have to make sure their messages lead to sufficient political action. The two groups have met together, and expect that, with few if any exceptions, their recommendations will be complementary, thus amplifying their effects. The Oceans Act also requires the Watkins Commission to present its report directly to the White House, and the government to act on it. The Pew Commission, by contrast, will have to fight for legislative attention with the wave of reports produced each year by non-profit organizations.

Representatives on the Oceans Caucus say they will help to publicize the commissions' findings and to formulate the legislation needed to enact them. There is bipartisan support in both the House and the Senate for many ocean issues, so the recommendations should receive significant attention. But the commissions know that not all of their results will necessarily be implemented. Molnia warns that the Watkins report may lose some of its impact because the commission was mandated by the Clinton administration, although its members were selected by Bush.

Despite these uncertainties, those involved are confident that the reports will be catalysts for change, even if the scale of that change remains unclear. "Whether we will go all the way to a Department of the Oceans I don't know," says Rufe. "But I hope we'll end up with something more substantial than just a rearrangement of the deckchairs."

Mark Schroppe is a science writer in Melbourne, Florida.

Stratton report

♦ www.lib.noaa.gov/edocs/stratton/title.html

Watkins Commission

♦ www.oceancommission.gov

Pew Oceans Commission

♦ www.pewoceans.org

NMFS report 2002

♦ www.nmfs.noaa.gov/sfa/reg_svcs/statusostocks/Stock_status01.htm

NOAA



A 'wet NASA': the Maryland home of the National Oceanic and Atmospheric Administration.

Caught on camera

By recording animal movements that are too fast for the human eye to follow, high-speed digital video is transforming studies at the interface of biomechanics, neuroscience and evolutionary biology. Rex Dalton reports.

One spring day last year, deep in the lush tropical cloud forest of Ecuador's Maquipucuna Reserve, Kimberly Bostwick positioned herself on the side of a ravine to observe an avian mating show in the upper canopy. Pulling out her new high-speed video camera, she waited for the brightly coloured actors: male club-winged manakins (*Machaeropterus deliciosus*).

Ornithologists knew that the male manakin makes a musical humming sound with its wings as part of its mating ritual¹, but no one had ever seen exactly how it was done. Bostwick's camera captured for the first time one of the birds clapping its wings behind its back at blinding speed. "It brought tears to my eyes," recalls Bostwick, now a curator of ornithology and mammalogy at Cornell University's Museum of Vertebrates in Ithaca, New York. "It was so gorgeous."

Frozen in time

Bostwick's video observations, contained within the doctoral thesis she completed at the University of Kansas in May, are part of a new wave of studies made possible by the relatively inexpensive, portable, high-speed cameras that have hit the market in the past few years. They have made it possible to take freeze-frame studies of lizards running on water, flies beating their wings, cockroaches scuttling over obstacles and fish gliding through water. By integrating this information with electrophysiological recordings,

researchers are gaining previously inaccessible insights into animal movement and its neural control. "We are creating a new field of neuromechanics," says Robert Full, who studies animal locomotion at the University of California, Berkeley.

Evolutionary biology is another beneficiary. For Bostwick and her Kansas supervisor, Rick Prum, the cameras have shed light on the importance and intricacies of sexual selection, the evolutionary mechanism involving competition between males and mate choice by females that Charles Darwin proposed can operate alongside natural selection.

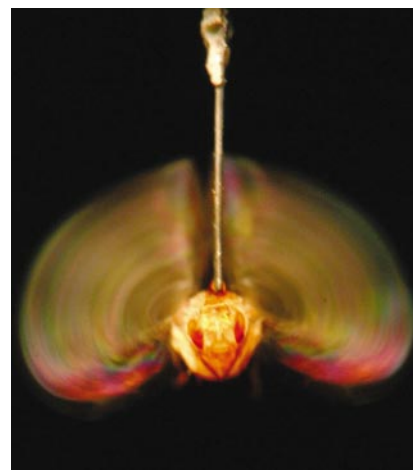
In 1871, Darwin wrote about sexual selection as a driving force for the evolution of mechanical sound production by birds². Manakins are the most extreme example, the males having evolved flattened wing bones and thick-tipped feathers to facilitate their sonic displays. Bostwick's video observations have now revealed that sexual selection has pushed the mechanics involved to the physiological limit: the birds' wings move back and forth almost as fast as muscle cells can contract. "The feathers hit behind the back 120 times per second," says Bostwick.

Rapid response

Just five years ago, a camera system to shoot this event at the required speed and light settings would have cost more than US\$30,000. "And they were almost impossible to work with in the field," says Bostwick. The camera she used, a MotionMeter manufactured by Redlake of San Diego, California, can shoot up to 1,000 frames per second. The hand-held device gives instant frame-by-frame playback and costs about \$10,000.

George Lauder, an evolutionary biologist now at Harvard University, was the first to explore the potential of high-speed cameras in mechanical studies of animal movement in the mid-1980s. Working at the University of California, Irvine, Lauder and his colleagues photographed fish swimming against a current in a flow tank, initially using old-fashioned film. As a result, their progress was limited by lengthy delays for film processing.

High-speed digital video has changed all



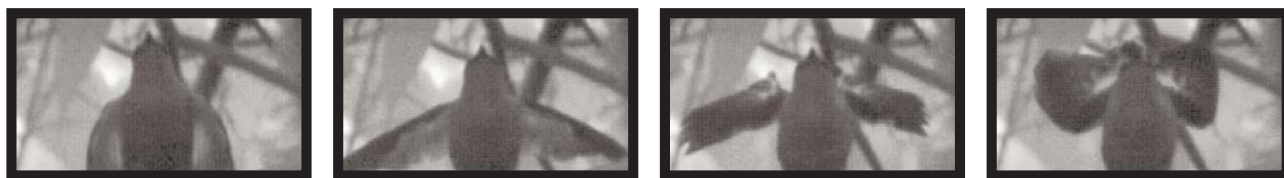
By suspending fruitflies on a wire, researchers can photograph their beating wings — but they remain a blur (above). High-speed video is now bringing this motion into focus.



In the frame: video images reveal how green basilisks manage to walk across water.

T. HSIEK

M. DICKINSON



which is revealed by the presence of tiny suspended glass beads that catch the light. Using this method, Lauder has, for instance, measured the forces involved in the doughnut-shaped vortices that form around the pectoral fins — behind the gills — of the bluegill sunfish (*Lepomis macrochirus*), in which the flow loops round before jetting through the doughnut's centre³.

By repeating their studies with fish from primitive sharks, through intermediate forms such as sturgeon, to the most advanced species, including the sunfish, Lauder and his colleagues hope to understand how complex mechanisms of aquatic locomotion have evolved and diversified.

Joint effort

Neuroscientist Roy Ritzmann of Case Western Reserve University in Cleveland, Ohio, is another enthusiast for high-speed video. He learned of its potential while serving on a panel reviewing one of Lauder's grant applications. Ritzmann now uses similar cameras to study deathhead cockroaches (*Blaberus discoidalis*), and has found that there is much more to the creatures' scuttling motion than meets the eye. "Each leg has a different function," says Ritzmann. "When the animal is moving, the joints are doing different things — legs are reaching out, pushing or pulling. It is really a complex situation."

Most recently, Ritzmann has watched roaches climb over obstacles in their path. The video stills reveal how the insects rotate the joints of their middle pair of legs to achieve the rearing necessary to clear the obstacle⁴, and electrophysiological recordings from the roaches' leg muscles have begun to reveal how this movement is controlled⁵. Ritzmann now plans to video cockroaches in which specific neural circuits have been disabled, to gain further insights into the processes involved.

Ritzmann's findings are fascinating from a purely biological standpoint, but they are also stimulating the field of robotics. His data are being studied by Roger Quinn and his colleagues at Case Western's Biologically Inspired Robotics Laboratory, who are building insect-like six-legged robots. For more than a decade, Quinn's group has tried to translate the roach's ability to traverse complex terrains to a mechanical device. "We are developing our fifth generation of robot — getting closer and closer to a cockroach," says Quinn.

Whereas Lauder and Ritzmann focus on



In a flap: Kimberly Bostwick has used high-speed video to show how male manakins clap their wings to generate a humming sound.

movement in water and on land, Michael Dickinson of the University of California, Berkeley, is using high-speed video to study flight. To analyse the flight of fruitflies and blowflies, arguably nature's most stunning flying machines, Dickinson and his colleagues have used a Kodak EktaPro camera that can shoot 3,000 frames a second.

Each cycle of wing movement has four components: upstroke, downstroke and two rotational phases, called pronation, which occurs after each upstroke, and supination, after each downstroke. By slowing down these movements using high-speed video, and then measuring the forces involved using the 'Robofly', a mechanical device that simulates a fly's wing movements in a vat of oil, Dickinson and his colleagues are uncovering the mysteries of insect aerodynamics.

Creating a buzz

They are currently trying to determine the forces involved when a fruitfly makes a rapid right-angle turn during flight, known as a saccade. A fly inside a cage is coaxed through a 'vision field' — a cube about seven millimetres square — where three high-speed video cameras capture its movement from different angles. The film is then slowed down, the wing and body movements are identified, and this knowledge is transferred to the Robofly, which then mimics a saccade and allows the forces involved to be measured. "Without high-speed video, we couldn't measure what the fly is doing," says Dickinson.

In other experiments, Dickinson and his colleagues have tethered blowflies (*Calliphora vicina*) with microscale electrodes implanted in 'steering' muscles that can subtly alter the plane and amplitude of each wing stroke.

They have found that the precise timing of spikes of activity in these muscles is the key to the flies' extreme manoeuvrability^{6,7}. "The video allows us to decode what the flight system is doing," says Dickinson, who will soon move to an expanded lab at the California Institute of Technology in Pasadena. "It has changed my view of biology."

In addition to bringing such fundamental insights, high-speed video is demystifying some of nature's quirks. Green basilisks (*Basiliscus plumifrons*), indigenous to Costa Rica, Nicaragua and Panama, are known as 'Jesus Christ lizards' for their ability to scamper across water. By shooting video at 250 frames per second, Tonia Hsieh, a graduate student in Lauder's lab, is investigating how they do it. "We are trying to figure out how locomotion changes on different surfaces," says Hsieh. "There is a huge body of literature on locomotion on the ground, but little on water and sand."

By videoing particles suspended in the water, Hsieh has examined the wake produced by the lizard's feet — which provided data on the vortices created, the force of impact of the feet on the surface and the forces exerted as the feet push backwards. At the Society for Integrative and Comparative Biology meeting in Anaheim, California, in January, she revealed that the forces that prevent the lizards sinking are generated by pushing the feet back and inwards, rather than back and outwards, like most lizards running on land.

This method allows young basilisks to stay above the surface — the largest that Hsieh has filmed running on water weighed 78 grams. But at 200 grams, adult basilisks no longer have their miraculous powers.

As more researchers begin to experiment with the new digital cameras, expect further wonders to emerge from nature's high-speed blur. Ritzmann likens the technology's current impact on organismal biology to that of electron microscopy in cell biology some three decades ago. "It allows you to see things never imagined before," he says. "It's almost like cheating."

Rex Dalton is Nature's US West Coast correspondent.

1. Prum, R. O. *Anim. Behav.* 55, 977–994 (1998).
2. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, London, 1871).
3. Drucker, E. G. & Lauder, G. V. *J. Exp. Biol.* 202, 2393–2412 (1999).
4. Watson, J. T., Ritzmann, R. E., Zill, S. N. & Pollack, A. J. *J. Comp. Physiol. A* 188, 39–53 (2002).
5. Watson, J. T., Ritzmann, R. E. & Pollack, A. J. *J. Comp. Physiol. A* 188, 55–69 (2002).
6. Tu, M. S. & Dickinson, M. H. *J. Comp. Physiol. A* 178, 813–830 (1996).
7. Balint, C. N. & Dickinson, M. H. *J. Exp. Biol.* 204, 4213–4226 (2001).

Plant mathematics and Fibonacci's flowers

Asymmetric cell division is an intriguing but unlikely explanation for the patterns of leaves.

Sir — The generation of leaf pattern (phyllotaxis) has long been a topic of interest and debate among plant biologists and mathematicians, as observed by Amar Klar in his Concepts essay¹. As Klar points out, the various models proposed for the production and maintenance of such biological patterns lack full supporting experimental data. These models, nonetheless, are based on (and reflect) biologically relevant observations and, contrary to Klar's assertion, lead to testable predictions. In contrast, Klar's new model based on asymmetric stem-cell division does not take into account some relevant biological observations that do not support his hypothesis.

During the maintenance of phyllotactic patterning, nascent multicellular organs (leaf primordia) determine the position of incipient primordia. For example, experimental data show that primordia can be initiated in the 'wrong' position by modifying either the hormonal² or biophysical^{3,4} context of the tissue, and that these induced primordia provide a feedback

loop to determine the position of the next leaf. Most evidence supports short-distance chemical signalling as the mechanism involved in designating the site of primordium initiation, with biophysical alteration in cell-wall extensibility as a key executor of the morphogenic programme initiated. The biochemical nature of the 'morphogen' is unknown, but progress is being made in its identification.

With respect to the initiation of phyllotactic pattern, the first primordia are generated from a multicellular ball, the plant embryo. Mathematical modelling convincingly demonstrates how an asymmetric pattern of a theoretical morphogen can be generated from an initially uniform field and how such asymmetry could lead to phyllotactic patterns⁵. Again, the biochemical identity of the proposed morphogen is unproven, but candidate molecules have been suggested⁶. In contrast, histological and clonal analysis of higher plant apical meristems has so far failed to reveal any consistent pattern of stem-cell division or

stem-cell lineage in relation to leaf formation⁷. A phyllotactic model using asymmetric cell division would predict the existence of such a prepattern.

Moreover, disruption of cellular patterning in the meristem does not lead to disruption of phyllotactic pattern⁸. It may be possible to imagine asymmetric division guiding leaf position in bryophytes (in which a single apical cell does undergo repeated asymmetric division to generate daughter cells that become incorporated into leaf-like organs), but in angiosperms (flowering plants), Klar's phyllotactic hypothesis represents a model too far.

Andrew J. Fleming

Institute of Plant Sciences, Swiss Federal Institute of Technology, Universitätsstrasse 2, CH-8092 Zürich, Switzerland

1. Klar, A. J. S. *Nature* **417**, 595 (2002).
2. Reinhardt, D. et al. *Plant Cell* **10**, 1427–1437 (1998).
3. Fleming, A. J. et al. *Science* **276**, 1415–1418 (1997).
4. Pien, S. et al. *Proc. Natl Acad. Sci. USA* **98**, 11812–11817 (2001).
5. Meinhardt, H. et al. in *Symmetry in Plants* (eds Jean, R. V. & Barabe, D.) 723–758 (World Scientific, Singapore, 1998).
6. Liu, C. M. et al. *Plant Cell* **5**, 621–630 (1993).
7. Furner, I. & Pumfrey, J. E. *Development* **115**, 755–764 (1992).
8. Wyrzykowska, J. et al. *Development* **129**, 957–964 (2002).

Macroecology is distinct from biogeography

Sir — According to H. J. Fisher in Correspondence¹, the scientific concerns of the recently emerged discipline of macroecology are a subset of those of biogeography, and hence the former is simply a sub-discipline of the latter. As evidence, Fisher observes that some macroecologists have written books on biogeography. We believe that his view is mistaken.

Biogeography is the science that attempts to document and understand spatial patterns of biodiversity². Macroecology is a way of studying relationships between organisms and their environment that involves characterizing and explaining statistical patterns of abundance, distribution and diversity — exploring the domain where ecology, biogeography, palaeontology and macroevolution come together³. Some overlap of interests between macroecology and biogeography has always been acknowledged (as has too, for example, between ecology and evolution), but in definitions that make clear their separate aims and identities.

Fisher's point that the same scientists have written on biogeography and macroecology in fact highlights precisely

the fact that they recognize the distinction.

Differences between macroecology and biogeography in scope and aim are apparent from the subjects covered in the macroecology symposium that prompted Fisher's comments. These included the ecological and evolutionary implications of the scaling of vascular networks; intraspecific optimization as a cause of interspecific allometry; the relationship between life history, population dynamics and extinction risk; neutral models of the assembly of local ecological communities from regional metacommunities; whether speciation rates are influenced by body size; and whether diversification is driven by key innovations (see ref. 4).

All these issues in macroecology may help us to understand the distribution and diversity of life on Earth, and so they may inform a range of disciplines that includes biogeography. But they are not biogeography.

Tim M. Blackburn*, Kevin J. Gaston†

**School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

†Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

1. Fisher, H. J. *Nature* **417**, 787 (2002).
2. Brown, J. H. & Lomolino, M. V. *Biogeography* (Sinauer Associates, Sunderland, Massachusetts, 1998).
3. Brown, J. H. *Macroecology* (Univ. Chicago Press, Chicago, 1995).
4. Nee, S. *Nature* **417**, 229–230 (2002).

The search for general principles in ecology

Sir — According to H. J. Fisher in Correspondence (*Nature* **417**, 787; 2002), macroecology is largely contained within biogeography, which documents and interprets patterns in biodiversity at large temporal and spatial scales. I disagree.

Macroecology aims to identify general principles or natural laws underlying the structure and function of ecological systems, which are apparent in the patterns of distribution and abundance of entities composing these systems, no matter what the scale of the analysis. Macroecology analyses some of the same patterns (for example, latitudinal patterns in species diversity), but its emphasis is not restricted to patterns apparent at large spatial scales, nor to contingent explanations.

Thus, macroecology can be understood as an approach to the study of ecological systems centred on the search for general and invariant principles underlying their diversity and variability. It is neither biogeography nor a large-scale version of community ecology, but a new overall perspective on ecological complexity.

Pablo A. Marquet

Centro de Estudios Avanzados en Ecología & Biodiversidad, Pontificia Universidad Católica de Chile, Casilla 114-D, Santiago, Chile

Encounters with a dark lady

An insight into one of the scientists who determined the structure of DNA.

Rosalind Franklin: The Dark Lady of DNA

by Brenda Maddox

HarperCollins: 2002. 380 pp. £20, \$29.95

David Blow

Rosalind Franklin, who provided the key data that suggested a double-helical structure for DNA, remains one of the most famous women scientists of her generation. Much of her fame derives from the obsessive behaviour of her collaborator/competitor Jim Watson. His accounts of his role in the discovery of the structure of DNA denigrate 'Rosy', who died in 1958. Had she survived until 1962, she might perhaps have shared the Nobel prize. As it is, she is better known to the general public than most Nobel laureates.

The insults to Franklin and others in Watson's first drafts of *The Double Helix* (Athenaeum, 1968) brought an angry response from Francis Crick, and Max Perutz declared that he was "furious". Linus Pauling even tried to veto its publication. Amidst these and many other objections, Watson made small changes and added a "pious" epilogue (Brenda Maddox's word). This admits: "My initial impressions of her (as recorded in the early pages of this book) were often wrong... The X-ray work... is increasingly regarded as superb... We both came to appreciate her personal honesty and generosity." In the face of powerful academic opposition, Harvard University Press refused to publish the work, forcing Watson to find another publisher.

In 1951, Maurice Wilkins, deputy director of the biophysics unit at King's College, London, obtained a sample of DNA that could be drawn into beautiful fibres. But Franklin, an externally funded research fellow administratively responsible to him, was the person with the technical insight and manipulative skill to produce superb X-ray-diffraction photographs from them. She discovered two states of the fibres, A and B, and learned how to prepare specimens purely in either form. The resulting photographs provided the underlying data from which the detailed helical organization of DNA was to be deduced.

Those who had even the slightest acquaintance with Franklin know that the criticisms of her that permeate *The Double Helix* do not reflect the real person. The misrepresentations provoked her friend Anne Sayre to write a detailed documentation of her life, character and research, *Rosalind Franklin and DNA* (Norton, 1968), which refutes, point by point, much of Watson's material. But the true circumstances of her



Egalité: Rosalind Franklin (left) enjoyed her time in Paris, where she got on well with her colleagues.

dreadful relationship with Wilkins while working at King's College have always been enigmatic, and this is the central theme of Maddox's book. Why did Wilkins and Franklin dislike each other so intensely? How did it come about that Franklin's spectacular X-ray-diffraction photographs of the A and B forms of DNA, which she freely shared with Pauling's collaborator Robert Corey, were not openly available to the Cambridge workers? Why was her crucial photograph of the B form, shown privately by Wilkins to Watson, not acknowledged in Watson and Crick's announcement of the structure of DNA?

There are other books that present the relevant facts, but Maddox paints a portrait with many more dimensions. Beginning with Rosalind's background in a long-established Jewish family, her father in a merchant bank and her uncle in the House of Lords, safely wealthy if not overwhelmingly rich, Maddox introduces a person who expects to be acknowledged and treated politely, who is assured enough to stand her ground and refuse to be browbeaten, and who becomes combative under pressure.

Aged 26, Franklin took a research appointment in Paris, where she enjoyed daily lunchtime discussions with her colleagues on politics, philosophy and human rights — "French intellectuals playing the part of French intellectuals". She went swimming and mountaineering, but in Paris she quickly adopted the fashionable New Look of 1947. She made good friends

among her male colleagues and her research progressed well. She did not mind dirty maintenance jobs in the lab, but dressed stylishly for important occasions. However, her aggressive intelligence had its harsh side, and she could easily be stung into defending territory that she saw as her own by means of vicious attack.

In early 1951 she returned to London with an independent fellowship at King's College. Maddox says that Franklin felt angry and excluded when she found that women were not allowed to lunch in the senior common room. Her male colleagues included ex-military men who had a tough, ferocious approach that spilled over into beer-drinking camaraderie, and were not used to treating women as their peers. Her results were considered Wilkins' property, obtained by her as a member of his team. Gentle feminine deference was not her style. She had experienced more balanced attitudes to gender in France, and she despised her seniors at King's for their manipulative manoeuvres and, above all, for their assumption that, as a woman, her role was a subordinate one. It is revealing how she blossomed later, when she moved to the more liberated atmosphere of Birkbeck College, less than a mile from King's.

Although Maddox helps us to appreciate how this civilized, idealistic and lively woman could truly be a menacing Dark Lady to Wilkins, this remains the greatest paradox. Franklin and Wilkins had both studied at Cambridge, and both had worked

New in paperback

Flutterland: Like Flatland, Only More So

by Ian Stewart
Perseus, \$14

"It is clear that Stewart knows the difference between serious science and solemn science. The book is as amusing and entertaining as it is informative, and serves as a fascinating introduction to the various mathematical spaces we inhabit." Lisa Lehrer Dive & Andrew Irvine, *Nature* **411**, 240–241 (2001).

The Cichlid Fishes: Nature's Grand Experiment in Evolution

by George W. Barlow
Perseus, \$18

"In this book, which is written in a colloquial and non-technical style, Barlow summarizes decades of his experience with cichlids... He draws on his impressive knowledge of the behaviour and natural history of many groups of fishes to make more general, important points about evolution... The book is entertaining and fun to read." Axel Meyer, *Nature* **410**, 17–18 (2001).

Intellectual Property Rights, Trade and Biodiversity

by Graham Dutfield
Earthscan, £24.95

Almost Human: A Journey into the World of Baboons

by Shirley C. Strum
University of Chicago Press, \$18

Out of its Mind: Psychiatry in Crisis — A Call for Reform

by J. Allan Hobson & Jonathan A. Leonard
Perseus, \$17

Atom: An Odyssey from the Big Bang to Life on Earth... and Beyond

by Lawrence M. Krauss
Abacus, £9.99

I Wish I'd Made You Angry Earlier: Essays on Science, Scientists and Humanity

by Max Perutz
Oxford University Press, £9.99

overseas. They could converse amicably about theatre, books and politics, but Wilkins found her peremptory, off-handed and 'spiky'. For whatever reason, they failed to communicate about science, and neither trusted the other.

At 23 years of age, the immature and intellectually overconfident Watson could not handle this situation. Wilkins showed him Franklin's results, but Watson was unable to deal straightforwardly with her — he says he was actually afraid of her, although Maddox thinks that "the male fear of the female has always been absurd". Assured of his own intellectual superiority, Watson believed that Franklin's approach was too rigorous, and that she lacked the mental adventurousness to transform the diffraction photographs into a structure. In fact, her insight led closer to the truth than Crick and Watson's first uninformed guess in November 1951, or Pauling and Corey's hasty proposal in early 1953.

Maddox's book opens a window into the minds of the central players. She has spent considerable effort interviewing the protagonists and everyone around them. Less partisan than Sayre, she avoids general philosophical questions by keeping more closely to the story. Her weakness is in presenting the scientific background. The oversimplified chapter introducing the role of DNA in biology contains confusing errors, and descriptions of diffraction techniques have distracting inaccuracies. She gives no insight into the differences between the A and B forms of DNA, the process of going from X-ray-diffraction observations

to a structural interpretation, or the significance of the twofold axis in DNA crystals.

These points can easily be ignored by a scientifically literate reader, and are perhaps irrelevant to the wider audience that the book will attract. For those who do not need the scientific issues explained, the book provides a thrilling commentary on the unusual characters of Franklin, Wilkins and Watson.

David Blow is in the Biophysics Group, Blackett Laboratory, Imperial College of Science, Technology and Medicine, London SW7 2BW, UK.

Record of a revolution

An Annotated Census of Copernicus' *De revolutionibus*: (Nuremberg, 1543 and Basel, 1566)

by Owen Gingerich
Brill: 2002. 404 pp. \$132

Adrian Johns

Arthur Koestler once described Copernicus' *On the Revolutions* as "the book that nobody read". Unlike landmark publications such as Darwin's *Origin of Species*, the volume that launched the astronomical revolution of the Renaissance had been "an all-time worst seller". In Koestler's view, the great miracle of that revolution was not so much that Copernicus destroyed geocentrism, but that anyone noticed.

Owen Gingerich's census of *De revolu-*

tionibus lays Koestler's verdict to rest, but in an intriguing and unexpected way. Gingerich has spent 30 years pursuing every single known copy. His has been a monumental enterprise, without parallel in the history of science. The result shows that, contra Koestler, Copernicus' book was not particularly rare — its print-run of 400–500 was quite usual for the period — and astronomers, at least, did indeed read it. The depredations of Catholic censorship, we can now tell, were real in Italy but virtually non-existent everywhere else. Most interestingly, however, Gingerich's survey begins to reveal how astronomers read the book — and how their readings coalesced into a copernican consensus.

Gingerich's careful recording of marginalia — including many photographs — permits us to see the readers of *De revolutionibus* emerging as a community. Notes and calculations originating with a handful of early devotees — Wittenberg professor Erasmus Reinhold, itinerant mathematician Paul Wittich, Tübingen scholar Michael Maestlin and Rhenish astronomer Jofrancus Offusius — were read, copied and recopied by successive generations of students. Their readings thus spread across Europe, establishing canonical methods, examples and critiques. Only the kind of detective work that Gingerich has now done could possibly have revealed these patterns of inheritance and emulation, out of which copernicanism itself grew.

This census is a great achievement. By



Read revolution: annotated copies of Copernicus' *De revolutionibus* reveal his readers' responses.

bringing us as close as we'll ever get to seeing Copernicus' readers at work, struggling to understand and use his demanding text, it reveals a scientific revolution in the making. ■

Adrian Johns is in the Department of History, University of Chicago, 1126 East 59th Street, Chicago, Illinois 60637, USA.

Eco-physiology comes of age

Physiological Ecology of Vertebrates: A View from Energetics

by Brian K. McNab

Comstock: 2002. 575 pp. \$75

John Speakman

Physiological ecology has been an emerging subject for the past 50 years, since its gestation in the seminal work of physiologists such as Knut Schmidt Nielsen, Per Scholander and Laurence Irving, who first looked at the physiology of animals in the wild. Brian McNab published his first paper in this area in 1963, and so has been on the scene since the field was in its infancy. He was involved in its development during the 1960s and 1970s, when it consisted of little more than compiling stories of wonderful animals living in harsh places. He played an active part in the sometimes heated debates during the difficult 'adolescent' years, when the field was trying to define itself — exemplified by *New Directions in Ecological Physiology* (Cambridge University Press, 1988), edited by Martin Feder *et al.* — and its final emergence as a mature field during the late 1990s.

The timing of *Physiological Ecology of Vertebrates* really could not have been better. McNab has apparently been on the cusp of writing it for the past 30 years, but I can see why he has never consolidated his ideas into a final volume until now. The field has really been in too great a flux over the past decade to allow any definitive summary to be produced, and before that it would have been little more than a catalogue with no unifying thread.

This is a book about physiology in wild vertebrates that contains no molecular biology. If we pursue the analogy between the development of the field and the development of a person, then this is a field in its confident early 20s, able to take stock of the issues that dogged its teenage past. This is the time to make a defining statement about where we are, because looming on the horizon is a mid-life crisis for this field, to do with how physiological ecology will accommodate the genomics revolution. So the timing seems appropriate, and I am glad McNab didn't wait another 10 or 15 years



When left high and dry, the tree frog *Phyllomedusa sauvagei* secretes a waxy coat to reduce water loss.

to write this book because by then it would have been unpublishable.

The book serves as a defining statement of where physiological ecology is right now. It is encyclopedic in its coverage, and there are almost as many references from the first 20 years of McNab's career as the past 20. I think that this will turn out to be a major strength of the book because for graduate students who think that science started in 1980 it will be a revelation to find that a lot of good science was done in the 1960s and 1970s that for all intents and purposes is lost because it is not included in modern computerized literature databases. If the book serves only to open the eyes of some students to this rich data field it will have achieved something. But I think this book will achieve much more than that.

It is a book that graduate students and established scientists alike will revel in. Undergraduates, too, will like the straightforward style of most of McNab's writing. Although it may be too expensive and probably a bit too detailed to work as an undergraduate course text, as a source of additional reading, library copies will be well used. I am certain that if I take this back to my office one of my postdocs or students will borrow it within days.

I enjoyed dipping into the book to read snippets about fields with which I am much less familiar than my own. Everyone will find favourite bits about which they were previously unaware. For me it was the 'waterproof frogs' that cover themselves with a layer of lipids but then have to remain motionless so that the waterproof veneer doesn't crack. Perhaps more importantly, I never felt that the coverage of my own field was inadequate.

McNab has a reputation among eco-physiologists for taking an unconventional stance on issues in the field. During the 1980s

and early 1990s he engaged in a protracted series of debates about the virtues of using phylogenetic contrasts to correct for the lack of independence in comparative studies. More recently he has been vocal on the merits and demerits of using mass-specific or non-mass-specific units to express basal metabolism. So I had expected a rather one-sided view of these issues to emerge in the book, but instead the treatment of these subjects is very well balanced — although McNab's own view is clear on most of the topics.

The book shows that physiological ecology has come of age. It contains over 3,000 references and, although it omits lots of individual studies, it still provides comprehensive overall cover. As a defining text covering 40–50 years of research it is unsurpassed and will become established as a defining reference.

Will this be the text that defines where the field goes in the coming decades, as promised in the foreword by James Brown? I actually don't think it will. The biggest thing that will happen in physiological ecology in the next 40 years is its integration with genomics, and the book doesn't touch on this issue at all. But to be honest, I don't think that matters. After all, Julian Huxley's *Evolution: The Modern Synthesis*, which came at a similar stage in the field of evolution, is still a classic. It doesn't really matter that because it was written in 1942, long before we knew the structure of DNA and the mechanisms by which the whole thing works, it didn't mention these topics. So the fact that McNab's book doesn't touch on genomics shouldn't stop it becoming a classic text in the field. ■

John Speakman is in the Department of Zoology, University of Aberdeen, Aberdeen AB24 3TZ, and the Division of Appetite and Energy Balance, Rowett Research Institute, Aberdeen AB21 9SB, UK.

Getting out of shape

Christopher M. Dobson

Two greatly debilitating illnesses — Alzheimer's disease, which is largely associated with ageing, and variant Creutzfeldt–Jakob disease (vCJD), which is linked with the recent bovine spongiform encephalopathy (BSE) epidemic in Britain — have had a huge impact on public consciousness. Despite their very different origins, these diseases are closely related to each other and also to the 20 or so other 'amyloidoses' — so called because they involve the aberrant deposition of proteins in the form of amyloid fibrils or plaques. Other diseases in this group, which are less publicized but no less devastating to those affected, are Parkinson's disease, type II (late-onset) diabetes and rare conditions such as familial insomnia.

Amyloidoses thus include some of the most feared and costly diseases in the Western world. Alzheimer's may very soon be the most prevalent and socially disruptive illness in the ageing populations of all developed countries. It is remarkable, therefore, that most of these diseases were virtually unknown until relatively recently. Indeed, the first detailed description of an amyloid pathology was made less than 100 years ago, when Alzheimer identified the form of dementia that bears his name. As with many

amyloidoses, Alzheimer's disease is usually sporadic, although there are less common hereditary forms that can afflict relatively young people — Alzheimer's most famous patient was in her early 50s.

The first form of a transmissible human amyloid disease was recognized only in the 1950s, when the fatal condition kuru, which was afflicting people in Papua New Guinea, was found to be the consequence of ritualistic cannibalism. Shortly afterwards, a similar 'transmissible spongiform encephalopathy' (TSE) emerged, this time in Europe and the United States, when a form of CJD struck individuals who as children had been treated with growth hormone extracted from human cadavers. And, of course, the consequences of the BSE catastrophe in the 1980s threaten many more people. Although a widespread epidemic of vCJD as a result of eating infected beef seems unlikely, about 100 people have already died of this truly dreadful illness. As other animal species besides cows and sheep are now known to be infected with similar conditions, including deer and elk in parts of North America, further tragedies may be waiting to unfold.

As mentioned above, all of these diseases, whether sporadic, familial or transmissible, are associated with deposits in tissue of proteins that are normally soluble. The deposits, depending on the disease, can be in the brain, in skeletal tissue or in other organs. The amount of protein involved ranges from scarcely detectable quantities to kilograms. Affected tissues are often found to be riddled with thread-like fibrils, sometimes assembled into plaques, with a single predominant protein component that is characteristic of each disease (for example, the prion protein in TSEs). This observation led to the idea that the structures of the soluble forms of the proteins involved have an unusual ability to convert to an alternative conformation, which can then assemble into the thread-like structures. A major puzzle was that although the proteins involved have different structures in their soluble forms, the fibrils look remarkably similar, suggesting that the molecular structures in the aggregated forms are essentially the same.

While investigations of the emerging amyloid diseases were taking place in medical schools, research into the nature of protein structures was developing in chemistry and physics departments. In the 1930s it was discovered that many proteins can exist in two forms: a globular, soluble state and a fibrous form, which is often produced at extreme temperatures or pH values. As the globular forms were usually those with biological activity, virtually all attention was focused on

Protein-misfolding diseases

Molecular evolution has not always kept pace with our aspirations for a longer and better life.

them. Indeed, the observation of alternative forms of proteins (except in the case of the prion protein) was not linked to amyloid diseases until very recently, following accidental discoveries that 'ordinary' proteins can form fibrils *in vitro* that are indistinguishable from those extracted from patients. Moreover, with sufficient patience and cunning, conditions could be found in which seemingly any protein could form amyloid fibrils. This observation suggested that the ability to form such fibrils is 'generic' to proteins, although the propensity to form such structures under given circumstances can vary greatly from one protein to another.

Why does this alternative form of protein structure exist? And why have 'new' diseases associated with it apparently emerged only recently? Both of these questions can be addressed by considering the nature of biological evolution. We each produce about 50,000 different proteins, but this number is only a minute fraction of the countless ways in which the 20 different amino acids can be linked together to give polypeptides of the length that is typical in living organisms (about 300 residues). Natural proteins are therefore a very special group of polypeptides with remarkable properties that have emerged over several billion years of evolution. One factor in increasing the competitive ability of an organism is the efficiency of its information transfer, which is largely achieved by rapid molecular diffusion. Cells are very densely packed with molecules that can communicate rapidly with each other. The fact that proteins do not normally interact inappropriately or even precipitate in such crowded environments is a triumph of evolutionary design.

The astonishing complexity and variety of natural proteins is determined by the unique way in which the various side chains of the constituent amino acids pack together. In amyloid fibrils, however, the structure of the polypeptide chain is dominated by hydrogen bonding between the atoms of the main chain that results in the formation of extended β -sheets, rather than by the specific interactions of the side chains that dictate the structures of globular proteins. In fact, the generic fibrillar forms of proteins can be regarded as the intrinsic 'polymer' structure

T. PHILLIPS, COURTESY OF THE NATIONAL PORTRAIT GALLERY, LONDON



The novelist and philosopher Iris Murdoch, one of the most famous victims of Alzheimer's.

of a polypeptide chain. After all, synthetic polymers such as nylon (a polypeptide without side chains) often form fibrillar structures. However, evolutionary processes have selected sequences of amino acids with the remarkable ability to form monomeric structures in which the main chain is folded in a unique way within the mass of close-packed side chains, preventing it from interacting with other molecules.

The existence of such structures has enabled proteins to develop a vast array of biological functions. But their ability to revert to the 'primordial' structure lingers, because the main chain of the polypeptide is preserved as a common feature of all natural proteins. To maintain proteins in their uniquely folded states, they must be kept in a carefully controlled environment. Under some circumstances, even partial unfolding will expose significant regions of the polypeptide chain to the outside world, allowing the protein to aggregate and convert into amyloid fibrils. Once formed, the strong hydrogen bonding between molecules can make this process effectively irreversible. This ability of proteins to change conformation is the essence of the amyloidoses; in these diseases, the proteins have converted into the 'primordial' amyloid structure rather than remaining in their evolved, globular states.

Although the manifestation of such a structural change differs between amyloid pathologies, it is a common factor in all of them. But why, in these diseases, does the normally tightly regulated environment within an organism fail to maintain a particular protein in its soluble state? One way is if a genetic mutation were to produce a variant that is less stable and less cooperative than the wild-type protein, and will therefore tend to unfold and aggregate more readily. This mechanism can explain many familial amyloidoses; indeed, the age of onset of some of these diseases correlates closely with the extent to which the familial mutation destabilizes the native protein. In other conditions, such as some dementias, mutations might increase the propensity of an incompletely folded protein to aggregate.

A unique characteristic of the prion group of amyloid diseases is their transmissibility between individuals (and in some cases, with much lower probability, between species). Transmission probably results from ingestion, from an external source, of prion proteins that have already aggregated. As with crystallization, the formation of amyloid fibrils is 'seeded' by pre-formed aggregates, a phenomenon that might also be responsible for the rapid progression of sporadic diseases such as Alzheimer's once the symptoms become evident. Such a seeding mechanism could explain transmission of TSEs by injection of contaminated growth hormone, by cannibalism or by eating infected beef.

In other cases, the physiological environ-

The ability of proteins to change conformation is the essence of the amyloidoses — in these diseases, the proteins have converted into the 'primordial' structure rather than remaining in their evolved states.

ment may not always be sufficiently well controlled to ensure that protein unfolding and aggregation are avoided. The loss of such control may be a characteristic of old age, and hence could explain why many amyloidoses are associated with ageing. It is interesting that for simple organisms such as bacteria, the lifespan of an individual is less than that of its molecules, as the latter are passed on through cell division. Perhaps to evolve long lifespans it is important to replace molecules before damage or aggregation can reap its toll — the lifespan of our proteins is rarely more than a few weeks. Yet, remarkably, some proteins in specialized organs, notably the eye lenses, last a lifetime. This astonishing fact shows that evolution can produce tremendously robust proteins for a particular function, although the occurrence of cataracts (which are themselves protein aggregates that are large enough to scatter light) shows that even the eye's lens proteins are not immortal.

So why have these diseases apparently become so much more prevalent in recent years? In my view, this must be at least partly due to societies changing more rapidly than molecular evolution can respond. The proteins that have emerged under evolutionary pressure are normally robust enough to resist reversion to aggregated states, and 'chaperone' proteins help still further to protect against such changes. In old age, however, we see the imperfections of these mechanisms. Proteins can be engineered to be more resistant to aggregation than our natural ones. But it is unlikely that there has been evolutionary pressure to do better than is necessary to pass on our genes and to protect our offspring until they reach maturity; indeed, perhaps evolutionary pressure opposes old age. The recent much longer average lifespans of people in the developed world mean that we are now entering uncharted evolutionary territory.

The rapidity of change in our societies relative to the rate of biological evolution is also revealed in conditions that are not so

directly associated with old age. Modern medical practices provide many examples of events not previously experienced in evolutionary biology. In the context of this article, it has been found that the release of a protein, β 2-microglobulin, during blood dialysis almost invariably results in amyloid deposits in skeletal tissue after several years. Fortunately, now that this problem has been recognized, new procedures and improvements to the design of dialysis membranes should alleviate this problem. Likewise, evolution has been powerless to deal with the effects of injecting contaminated natural growth hormone — now thankfully avoided by use of recombinant protein.

Such an argument may also explain the origins of kuru and vCJD. Cannibalism is an activity that has presumably been eliminated largely as a result of social pressures acting during evolution. BSE has almost undoubtedly resulted from the highly unnatural practice of feeding young cows on the remains of old ones, with the disease then being transmitted to humans as vCJD. Both kuru and BSE have virtually disappeared as a result of effective action taken once their origins were understood.

What can we learn from this complex story? The discovery of the underlying origin of amyloidoses is tremendously important in view of the increasing impact of these diseases. It allows us the opportunity to develop new strategies for drug therapy, improved medical procedures and higher food standards. But the most fundamental message is that evolution acts slowly, and we are an impatient species. The properties of biological molecules have developed slowly in response to the pressures experienced by their host species. We are now in effective control of both our own evolution and that of many other organisms on the planet. We must therefore strive to understand completely the underlying principles of 'molecular evolution' to take advantage of the tremendous opportunities of this 'post-evolutionary' era. In this way, not only can we avoid accidentally unleashing new scourges on our descendants, but we can take advantage of our new-found knowledge to extend and enhance the quality of all our lives. ■

Christopher M. Dobson is in the Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

FURTHER READING

- Dobson, C. M., Ellis, R. J. & Fersht, A. R. (eds) *Phil. Trans. R. Soc. Lond. B* **335**, 129–227 (2001).
- Solomon, B., Taraboulos, A. & Katchalski-Katzir, E. (eds) *Conformational Diseases: A Compendium* (Karger, Tunbridge Wells, UK, 2001).
- Csermely, P. *Trends Genom.* **17**, 701–704 (2001).
- Ellis, R. J. & Pinheiro, T. J. T. *Nature* **416**, 483–484 (2002).
- Iverson, L. *Nature* **417**, 231–233 (2002).
- Ferguson, N. M. *et al. Nature* **415**, 420–423 (2002).

Mass tool for diagnosis

Matthias Mann

Efficient and sensitive methods to determine whether, and to what extent, a person is infected with malaria should help to improve treatment. A high-tech approach, using mass spectrometry, may be the answer.

Years ago my PhD adviser, John Fenn of Yale University, told me that he wanted to develop a mass spectrometer — the well-known tool for identifying and quantifying compounds and for determining their structure — that could be used in the doctor's office. People would deposit their sample in one end of the machine, and out from the other end would come a list of diagnoses based on unambiguous 'digital' mass-spectrometry data. Of course, this didn't actually happen; instead, Fenn invented electrospray mass spectrometry¹, which, together with matrix-assisted laser-desorption (MALDI) mass spectrometry², laid the foundation for the current boom in mass-spectrometry-based proteomics, one of the hottest fields in functional genomics³.

Out of the limelight, however, mass spectrometrists are busily inventing technologies that bring us ever closer to the routine use of mass spectrometry in diagnosis. In a paper just out in *Analytical Chemistry*, for instance, Demirev and colleagues⁴ describe how they devised an ingenious method to detect the presence of the malaria parasite at a sensitivity of just ten parasites per microlitre of human blood. The method, which was demonstrated with cultured parasites, addresses one of the most serious worldwide health problems⁵. It is clinically important to determine not only whether patients are infected with the malaria parasite, but also how heavily they are infected; that makes it easier, for example, to determine how effective a treatment is.

Demirev and colleagues' approach⁴ takes advantage of a peculiarity in the parasite's life cycle: while circulating in the human blood stream, the parasite grows inside red blood cells (Fig. 1). It builds up elaborate structures, including inner membranes, inside these cells, before bursting them to release progeny parasites. The material for this build-up comes almost entirely from the breakdown of haemoglobin, which constitutes 95% of the red blood cell. After haemoglobin has been digested by the parasite, the oxygen-carrying haem group remains and the parasite stores it in an almost crystalline form in special vacuoles. The haem group is photoactive and turns out to be easily detectable by direct laser-desorption mass spectrometry. As it builds up to very high levels inside the parasite,

it is a promising diagnostic marker.

In their technique, Demirev *et al.* exploit the fact that the membranes of red blood cells contain cholesterol, so as to lyse the cells while keeping the inner, cholesterol-lacking, membranes of the parasite intact. They then isolate the parasite and apply its contents to a metal target, on which many samples can be arrayed. The target is transferred into the vacuum system of a MALDI instrument and illuminated by short laser pulses, which volatilize and ionize the haem (assisted not by a matrix as in normal MALDI, but by the photoactive properties of the haem). The haem ions are then accelerated and shot down a flight tube under vacuum, and their mass is determined by the arrival time relative to the laser pulse — a technique called time-of-flight mass spectrometry. Because of the high concentration of haem in the parasites, and because the normal red-blood-cell contents have been removed, excellent sensitivity and discrimination between infected and uninfected blood can be obtained.

The authors show that as little as one parasite per microlitre should be detectable with more specialized equipment and optimized procedures. Moreover, characteristic fragmentation of the haem is observed under certain conditions, increasing the specificity of the measurement. The method also seems to be quantitative — traditionally a weak point with mass spectrometry, especially MALDI mass spectrometry, because of the many factors that influence peak height in a mass spectrum. Even if the quantity of haem could not robustly be determined from a normal mass spectrum, there are now many variants of stable-isotope methods that could provide effective quantification by simple matching of peaks to an internal standard.

How does this new method stack up against existing procedures? That depends on which parameter we look at. It seems clear that measuring the haem within parasites in blood is a clever way to diagnose malaria, and the method already compares well with current techniques in terms of sensitivity. History shows that once it gets a foothold into an analytical field, mass spectrometry usually conquers it — witness such disparate recent developments as the shift to the use of mass spectrometry rather than two-dimensional gel electrophoresis in proteomics, and

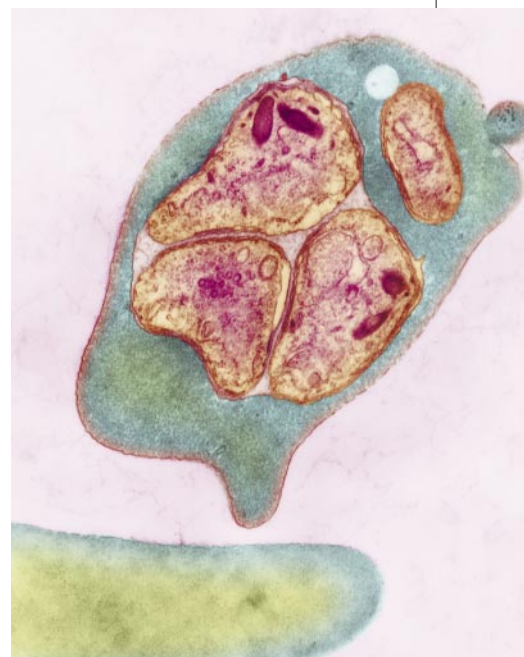


Figure 1 In the blood — the malaria parasite (brown) in a sectioned human red blood cell (green). Demirev *et al.* have developed a mass-spectrometry method to detect malaria parasites in human blood.

the routine analysis of single-nucleotide repeats in nucleic acids. Mass spectrometry's advantages are its precise, unambiguous and machine-readable output and the fact that measurements can be performed in a very short time if there are standard conditions. For malaria diagnosis, many samples could be prepared in parallel and measurement per sample should take no longer than a second or so.

However, malaria diagnosis is most needed in remote rural areas, often with no electricity. Today's 'gold standard' for diagnosis is the thick and thin blood smear, which consists of counting parasites in stained blood smears by microscopy. That method is decidedly low-tech, having been developed a hundred years ago⁶; it is difficult to perform; and in practice its sensitivity is often less than the World Health Organization's recommendation of 100 parasites per microlitre of blood. But it does have the great advantage of being deployable in the places where malaria is endemic. These are not hospitable environments for today's high-tech mass

spectrometers. That said, different types of mass spectrometer have been used in the field for more than ten years. For instance, battery-operated ones are used by the armed forces and in environmental work. Moreover, small mass spectrometers can operate in challenging environments (one was delivered to Mars by the Viking space probe), and even the time-of-flight spectrometer used with MALDI can in principle be miniaturized tremendously, possibly even to palm size⁷.

The biochemical purification steps of Demirev and colleagues' technique, including centrifugation, would also have to be compatible with tough environments. So it is interesting that analysis of whole blood by direct laser desorption does not result in an appreciable haem signal, presumably because the haem in normal red blood cells is tightly bound to haemoglobin (P. A. Demirev, personal communication). This suggests that much simpler preparation methods, without centrifugation steps to remove normal red blood cells, might be feasible. Finally, the method as described does not distinguish between different types of malaria parasite. But this could in principle be done by proteomics technology using mass spectrometry.

So the technical potential for a routine and relevant malaria-detection assay is certainly here, and further developments will — as usual — depend to a certain extent on the political will and economic or philanthropic interest in better methods to diagnose this predominantly tropical disease. We can hope that the current push to detect biological-warfare agents, in which mass spectrometry has a prominent role, will aid diagnostic efforts as a side effect. There are also many, equally promising, efforts to detect proteins that are characteristic of other diseases in blood or other bodily fluids — all fuelled by the rapid technical developments in mass spectrometry. Perhaps we are coming full circle to realize Fenn's dream of the mass spectrometer as a universal diagnostic machine.

Matthias Mann is in the Protein Interaction Laboratory, Department of Biochemistry and Molecular Biology, University of Southern Denmark—Odense, Campusvej 55, DK-5230 Odense M, Denmark.
e-mail: mann@bmb.sdu.dk

1. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. *Science* **246**, 64–71 (1989).
2. Hillenkamp, F., Karas, M., Beavis, R. C. & Chait, B. T. *Anal. Chem.* **63**, 1193A–1202A (1991).
3. Mann, M., Hendrickson, R. C. & Pandey, A. *Annu. Rev. Biochem.* **70**, 437–473 (2001).
4. Demirev, P. A. *et al. Anal. Chem.* **74**, 3262–3266 (2002).
5. Nature Insight: Malaria. *Nature* **415**, 669–715 (2002).
6. Moody, A. *Clin. Microbiol. Rev.* **15**, 66–78 (2002).
7. Cotter, R. J., Fancher, C. & Cornish, T. J. *J. Mass Spectrom.* **34**, 1368–1372 (1999).

Cell motility

Braking WAVES

Giles O. C. Cory and Anne J. Ridley

When cells move, they alter their internal skeleton to push membrane out at the front and pull it in at the back. New work fills in some of the gaps in our knowledge of how this process is regulated.

Cells come in many different shapes, depending on their function. Neurons, for example, extend long, branching protrusions (axons) to form the web of cellular connections found in the brain. Many other types of cell use changes in their shape to move, and this is an essential property of immune cells, as they crawl through tissues in their search for pathogens. Günther Gerisch's spectacular time-lapse film¹ of one such immune cell, a neutrophil, chasing a bacterium highlights the ability of cells to sense external signals and to direct their migration towards them. Not surprisingly, however, cell migration also has its dark side and, when not tightly controlled, can lead to cancer cells invading nearby tissues and spreading through the body. It has long been known that a small protein called Rac acts as a molecular switch that affects many aspects of cell shape and migration². On page 790 of this issue, Eden and co-workers³ shed light on a new mechanism by which Rac acts

upon another key protein in the movement business — WAVE1.

The shape and movement of our cells are controlled by their internal framework, the cytoskeleton. In particular, the highly dynamic and reversible polymerization of a cytoskeletal protein called actin plays a central role in regulating cell behaviour. Actin is soluble when in its monomeric, globular form, but upon polymerization forms insoluble filaments with significant mechanical strength. Localized actin polymerization at the front of a cell pushes its membrane forward in a sheet-like structure known as a lamellipodium². This can lead to the formation of cell extensions and, when coordinated with contraction at the rear, can propel cells at speeds of up to 30 micrometres a minute.

The rate-limiting step of actin polymerization is the assembly of three actin monomers into a trimer⁴. So polymerization is enhanced by several factors that stimulate

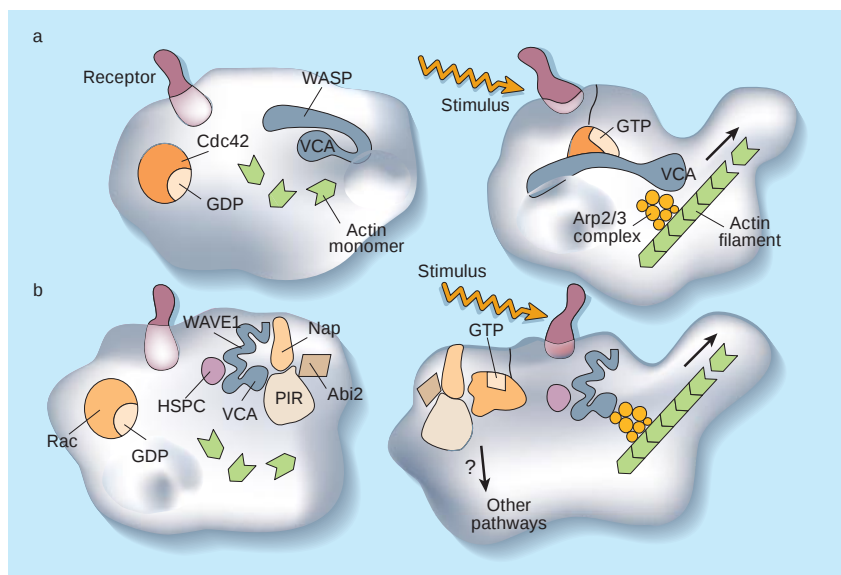


Figure 1 Regulation of actin polymerization, and hence cell movement, by the WASP and WAVE proteins. a, Left, in the absence of a motility stimulus, WASP is in an autoinhibited state. Right, extracellular signals lead to the activation of Cdc42 (it binds GTP instead of GDP), which then binds to WASP, exposing its VCA domain. This domain can now bind to and activate the Arp2/3 complex, which induces the assembly of actin into filaments that push the cell membrane forward. b, Left, Eden *et al.*³ have found that WAVE1 forms a complex with Nap, HSPC, PIR and Abi2 proteins, keeping WAVE1 inactive. Right, receptor stimulation leads to the activation of Rac (or recruitment of the Nck protein; not shown), which binds to the Nap–PIR–Abi2 subcomplex and liberates the HSPC–WAVE1 subcomplex, which again binds the Arp2/3 complex and stimulates actin polymerization. The remaining part of the complex may act upon other cellular pathways or perhaps influence the shape of the actin structure formed.

such 'nucleation'. A family of regulatory proteins, including the Wiskott–Aldrich syndrome protein (WASP) and WAVES 1, 2 and 3, achieve this by interacting with a complex of seven other proteins — the Arp2/3 complex⁴. When bound by a conserved region within WASP or WAVE (the so-called VCA domain), the Arp2/3 complex is activated and catalyses actin polymerization (Fig. 1), presumably by mimicking an actin trimer. WAVE1 is located at the edges of growing lamellipodia^{5,6}, and drives localized activation of the Arp2/3 complex and actin polymerization at the front of the protrusion.

WASPs and WAVES are themselves regulated by two other proteins², Cdc42 and Rac. These are part of a superfamily of proteins that act as specific molecular switches, controlling a plethora of cellular pathways from membrane trafficking to cell growth and division. They are activated by exchanging the small molecule GDP for GTP, and can then interact with certain downstream proteins. So, when active Cdc42 binds to WASP, it disrupts WASP's self-inhibiting structure and exposes the VCA domain, which can thereby activate the Arp2/3 complex (Fig. 1a)⁴.

In contrast, the WAVE proteins do not contain binding sites for Rac or Cdc42, and so how their activity is controlled has been a mystery. Two years ago it was shown that active Rac binds an intermediary protein (IRSp53), which in turn binds to WAVE2 to activate it⁷. But this intermediary does not bind WAVES 1 and 3. Furthermore, it appears that purified WAVE1 is constitutively active³, suggesting that it must somehow be repressed in areas of the cell that are not actively protruding.

Eden and colleagues³ now describe a novel control mechanism whereby WAVE1 is kept inactive, through its association with four other proteins: Nap125, PIR121, Abi2 and the tiny HSPC300. The authors show that this complex is unable to stimulate actin polymerization in an *in vitro* assay but that the addition of purified, active Rac relieves the inhibition. It does this by binding to the Nap125–PIR121–Abi2 subcomplex and dissociating it from HSPC300–WAVE1, which can then go on to nucleate actin polymerization (Fig. 1b).

Cdc42 and Rac are not the only keys able to unlock the activity of the WASP and WAVE proteins⁴. For example, WASP can be activated by binding to SH3 domains⁸ — modular protein domains found in several signalling molecules — independently of Cdc42. Similarly, Eden *et al.* show that Nck, a small adaptor protein, can also bind to Nap125–PIR121–Abi2 and induce the dissociation of HSPC300–WAVE1, independently of Rac. This adaptor also binds to certain receptors on the cell surface, and so may direct the spatial control of WAVE

activation and actin polymerization to sites of receptor engagement. Support for this idea comes from studies of enteropathogenic *Escherichia coli* bacteria, which bind to the outside of cells and recruit cellular Nck (presumably by mimicking a cellular receptor); this in turn causes localized actin polymerization⁹. Nck also binds to another of Rac's targets, the protein kinase PAK¹⁰, and it will be important to work out the relative contributions of Rac and Nck in regulating WAVES, PAKs and probably other proteins.

Interestingly, Nap125 belongs to a family of transmembrane proteins (the HEM family) that is found in organisms ranging from worms to humans, and it has been implicated in embryo development (itself a process requiring extensive cell movements)¹¹, axon growth¹² and the prevention of cell suicide¹³. Proteins from this family are found in a variety of tissues, including brain and immune cells. Given the results of Eden *et al.*³, it is possible that these proteins represent an evolutionarily conserved mechanism for regulating cell shape via WAVES. Abi2 may be important for localizing WAVES, as it is found at the front of lamellipodia¹⁴.

It remains to be seen whether the subcomplex of Nap125, PIR121 and Abi2, bound to Rac or Nck, has a signalling role in its own right, and how interfering with the function of this subcomplex affects cell

shape and movement *in vitro* and *in vivo*. No doubt there are still more ways to regulate WAVES, and another question is whether this subcomplex can bind to WAVES 2 and 3, or whether each WAVE is regulated in a unique way. In any case, these new results³ reveal the complexity of the signalling network surrounding WASP and WAVE proteins, reflecting their role in controlling the multitude of cellular functions governed by the actin cytoskeleton. ■

Giles O. C. Cory and Anne J. Ridley are at the Ludwig Institute for Cancer Research, Royal Free and University College School of Medicine, 91 Riding House Street, London W1W 7BS, UK.

e-mail: anne@ludwig.ucl.ac.uk

1. <http://www.hopkinsmedicine.org/cellbio/devreotes/neutrophil.mov>
2. Ridley, A. J. *Trends Cell Biol.* **11**, 471–477 (2001).
3. Eden, S., Rohatgi, R., Podtelejnikov, A. V., Mann, M. & Kirschner, M. W. *Nature* **418**, 790–793 (2002).
4. Higgs, H. N. & Pollard, T. D. *Annu. Rev. Biochem.* **70**, 649–676 (2001).
5. Hahne, P., Sechi, A., Benesch, S. & Small, J. V. *FEBS Lett.* **492**, 215–220 (2001).
6. Nakagawa, H. *et al.* *J. Cell Sci.* **114**, 1555–1565 (2001).
7. Miki, H., Yamaguchi, H., Suetsugu, S. & Takenawa, T. *Nature* **408**, 732–735 (2000).
8. Fukuoaka, M. *et al.* *J. Cell Biol.* **152**, 471–482 (2001).
9. Gruenheid, S. *et al.* *Nature Cell Biol.* **3**, 856–859 (2001).
10. Li, W., Fan, J. & Woodley, D. T. *Oncogene* **20**, 6403–6417 (2001).
11. Baumgartner, S. *et al.* *J. Mol. Biol.* **251**, 41–49 (1995).
12. Hummel, T., Leifker, K. & Klammt, C. *Genes Dev.* **14**, 863–873 (2000).
13. Suzuki, T. *et al.* *Genomics* **63**, 246–254 (2000).
14. Stradal, T. *et al.* *Curr. Biol.* **11**, 891–895 (2001).

Superconductivity

Mind the double gap

Warren Pickett

Magnesium diboride superconducts at an unexpectedly high temperature. It is now clear that the material also has an unusual, but long-sought, 'double energy gap' structure that influences its superconductivity.

In 1986, an 'earthquake' hit the field of high-temperature superconductivity: layered copper oxides with complicated crystal structures were discovered to superconduct at amazingly high temperatures, 90–100 K and above (in fact, the superconductivity record is now closer to room temperature than to absolute zero). Then came the aftershock, with the announcement by Jun Akimitsu¹ last year that a simple intermetallic compound superconducts at almost double the temperature of previously studied intermetallic compounds. Although not competing with the copper oxides, magnesium diboride (MgB₂) has a transition temperature of 39 K, and good prospects that uses may be found for it².

Our understanding of why this compound should superconduct at such a relatively high temperature has progressed rapidly in the past 18 months. Now Choi *et*

*al.*³ (page 758 of this issue) have calculated the superconducting structure of MgB₂ from basic atomic data and physical laws, and find that it is intrinsically a 'two-band' superconductor of a kind not previously seen. They have used computational techniques not only to obtain the correct value for the transition temperature of MgB₂, but also to map its two-band structure in detail.

By the early 1980s, superconductivity was thought to be well understood. 'High' transition temperatures (at that time, only around 20 K) required three conditions: a high density of charge carriers; electrons residing in open *d*-shell orbitals of the constituent atoms; and crystals possessing a cubic symmetric structure. Copper-oxide high-temperature superconductors don't fit this mould, but they are simply a different animal. Although it is a simple intermetallic compound, MgB₂ achieves its impressively

high transition temperature by breaking the rules.

Conventional materials superconduct when electrons form pairs ('Cooper pairs'). In MgB_2 , the pairing arises from attractive interactions between the electrons that are mediated by atomic vibrations, or phonons. This has been verified by isotopic replacement experiments and explained in some detail by several computational groups. But unlike previously known good superconductors, MgB_2 does not have a high charge-carrier density; only simple *s*- and *p*-shell electrons are involved in its superconductivity; and it has a non-cubic crystal structure. Other features make MgB_2 even more remarkable: only a few of the atomic-vibration modes are involved in its ability to superconduct, and only about half of its charge carriers are strongly affected during superconductivity.

Polycrystalline samples of MgB_2 can be prepared fairly easily, but the lack of single-crystal samples — which have been made under pressure only recently — has delayed definitive experimental results. However, decades of computational research into materials and the fact that MgB_2 is, both structurally and electronically, a simple material, have led to many of the questions about its superconductivity being addressed and answered by computational means. The results of Choi *et al.*³ are the most detailed to date, and are the first to reveal the full structure of a fundamental feature of superconductors, the energy gap.

The energy gap is the binding energy that induces electrons to form pairs, and it is directly related to the material's transition temperature. It determines the thermodynamic behaviour of the superconductor and influences its properties when used in devices (for example, pair tunnelling in Josephson junctions). In most known superconductors, the size of the energy gap varies only slightly for the various electron pairs involved in the superconductivity — it is considered to be isotropic. But for MgB_2 , the energy gap seems to be anisotropic — it has two different bands, that is, two different binding energies for electron-pair formation. This feature has been suggested by several groups and calculated by Liu *et al.*⁴ and by Golubov *et al.*⁵ in a simplified form. Choi *et al.*³ now confirm this picture in detail.

The two-band character of MgB_2 means that we must adjust our ideas on superconductivity. According to the usual picture, the main characteristics of a superconductor — its critical temperature and energy gap — are determined by a coupling strength, a repulsion strength and one or two representative vibrational frequencies. For a two-band superconductor the picture is much richer: there are four coupling strengths and four analogous repulsion strengths; but the average coupling strength is insufficient to deter-

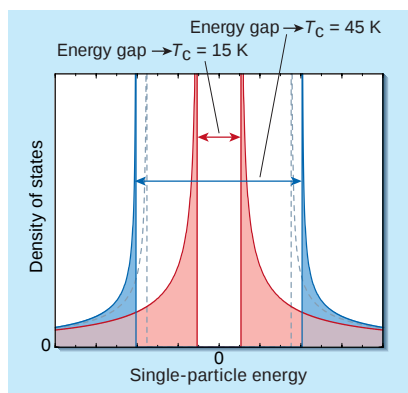


Figure 1 Double-gap superconductor. The 'energy gap' in a material reflects how easily its electrons form superconducting pairs, and so is related to the transition temperature up to which the material is superconducting. Conventionally, superconductors have a single energy gap in the spectrum of unpaired-electron states, for example the single gap indicated by the dashed lines here. The states in the normal (non-superconducting) phase are 'pushed out' of the gap region, leaving a sharp peak on either side of the gap. Choi *et al.*³ demonstrate, however, that magnesium diboride has two energy gaps, one corresponding to a transition temperature (T_c) of 45 K (blue), the other to 15 K (red). The result is a compromise transition temperature for the bulk material of 39 K — an unusually high value for a material that is not a high-temperature, copper-oxide superconductor.

mine either the transition temperature or the thermodynamic characteristics. The material's properties instead depend on two energy gaps (Fig. 1), one that would correspond to a transition temperature of 45 K, the other to a transition temperature of 15 K. Operating in concert, they lead to a transition temperature of 39 K. These calculated values^{3,5} are

Signal transduction

Positive feedback from coffee

Jean-Marie Vagueois

Caffeine acts on our nerve cells to wake us up. It turns out that it does so through a molecular signalling pathway that involves a positive feedback loop, boosting caffeine's effects from inside the cell.

Have you ever wondered why coffee is such an effective pick-me-up? The answer has been gradually coming into focus over the past few years, thanks to numerous studies of the brain regions and neuronal molecules affected. With the work reported by Lindskog and colleagues¹ on page 774 of this issue, much of the fine detail has been filled in.

In broad terms, the effects of caffeine on the body have been known for several years.

probably more accurate than, although consistent with, estimates from thermodynamic and spectroscopic experiments.

The two-band model of superconductivity has been a theorists' pet for four decades, but until now has remained just that. Kondo⁶ suggested two-band superconductivity driven by electron–electron repulsion which would be possible if the two gaps have opposite signs. But there is no evidence that this picture is relevant to MgB_2 . On the other hand, Leggett⁷ predicted that a two-band character would lead to a quantum-mechanically induced oscillation between electron pairs that possess different binding energies corresponding to the two gaps. The possibility of observing this effect in MgB_2 is being assessed⁸.

It is clear that MgB_2 is a new breed of superconductor, and other intermetallic compounds might also superconduct in the same way⁹. Another new breed is the small class of ferromagnets that show superconductivity¹⁰ — a property long thought impossible. It seems that all we can really count on is that new superconductors will continue to appear in unexpected ways. ■

Warren Pickett is in the Department of Physics, University of California, Davis, California 95616, USA.

e-mail: pickett@physics.ucdavis.edu

1. Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. *Nature* **410**, 63–64 (2001).
2. Grant, P. *Nature* **411**, 532–533 (2001).
3. Choi, E. J., Roundy, D., Sun, H., Cohen, M. L. & Louie, S. G. *Nature* **418**, 758–760 (2002).
4. Liu, A. Y., Mazin, I. I. & Kortus, J. *Phys. Rev. Lett.* **87**, 087005 (2001).
5. Golubov, A. A. *et al.* *J. Phys. Cond. Mater.* **14**, 1353–1360 (2002).
6. Kondo, J. *Prog. Theor. Phys.* **29**, 1–9 (1963).
7. Leggett, A. J. *Prog. Theor. Phys.* **36**, 901–930 (1966).
8. Sharapov, S. G., Gusynin, V. P. & Beck, H. Preprint cond-mat/0205131 (2002); <http://xxx.lanl.gov>
9. Rosner, H., Kitaigorodsky, A. & Pickett, W. E. *Phys. Rev. Lett.* **88**, 127001 (2002).
10. Saxena, S. S. *et al.* *Nature* **406**, 587–592 (2000).

1995, when Svenningsson, Nomikos and Fredholm³ found that the stimulatory effects of caffeine are strongly correlated with the reduced expression of certain genes that would normally be switched on in response to activation of the A_{2A} receptor by adenosine. More support came from studies of mice engineered to lack this receptor⁴. Moreover, A_{2A} -receptor blockers are being developed as potential treatments for Parkinson's disease, and there is evidence that caffeine can help to protect against the nerve degeneration associated with this disease⁵. So the idea that caffeine works by blocking the A_{2A} receptors seems well established.

But how does this lead to a stimulant effect? To understand this, we need first to look at the effects of adenosine on neurons (Fig. 1a). As mentioned above, adenosine is a type of neuromodulator, a molecule that fine-tunes the effects of the chemicals — neurotransmitters — that nerve cells use to communicate. Neuromodulators often do this by binding to receptors that are in turn coupled to common molecular switches, called G proteins, within neurons. Both neurotransmitters and neuromodulators produce many of their effects through complex, slow signalling pathways that modify other neuronal proteins by the addition and removal of phosphate groups (phosphorylation and dephosphorylation, respectively)⁶. This can change the physiological properties of the target proteins, which might be other receptors, ion channels or pumps, or gene-transcription factors.

For instance, by binding to and activating the G-protein-coupled A_1 or A_{2A} receptors, adenosine can control the enzyme adenylyl cyclase in opposite ways, with the A_{2A} receptors stimulating the enzyme and thereby leading to production of the messenger molecule cyclic AMP (cAMP). This eventually leads, via a cascade of protein modification, to the increased transcription of certain genes, to an expression level that is correlated with reduced motor activity. One protein cascade involves the enzyme protein kinase A, which is activated by cAMP and which in turn phosphorylates the DARPP-32 protein in a subgroup of neurons in the striatum⁷ — a brain region that controls motor activity.

Lindskog *et al.*¹ now provide compelling evidence for a role of DARPP-32 in mediating the effects of caffeine (Fig. 1b). The authors found that A_{2A} receptors are present in about half of the striatal neurons that contain DARPP-32. They next showed that the motor effects of caffeine depend on DARPP-32, as the effects were not seen in mice lacking this protein. Interestingly, however, this dependence could be overcome by increasing the dose of caffeine.

How can these effects of DARPP-32 be explained? Protein kinase A phosphorylates this protein at the amino acid threonine at position 34. That turns DARPP-32 into an

inhibitor of a certain protein-dephosphorylating enzyme (a protein phosphatase). But when another amino acid in DARPP-32, the threonine at position 75, is phosphorylated, the protein becomes an inhibitor of protein kinase A itself (Fig. 1b). So does caffeine lead to an increase in phosphorylation of threonine 75? Lindskog *et al.* show that it does, as does a drug that specifically blocks the A_{2A} receptor. And how exactly does this occur? The authors find that caffeine does not increase the constitutive activity of the neuronal Cdk5, an enzyme that phosphorylates DARPP-32 at threonine 75. This result and other data suggest that caffeine instead inhibits protein phosphatase-2A, preventing this enzyme from dephosphorylating threonine 75. That results in a higher level of DARPP-32 that is phosphorylated at this amino acid.

All of which leads to the following model. With low doses of caffeine, only a little cAMP is produced by adenylyl cyclase; protein kinase A is activated only weakly and cannot activate protein phosphatase-2A. So threonine 75 on DARPP-32 remains phosphorylated, and this protein in turn further inhibits protein kinase A, forming a positive

feedback loop. Such inhibition leads to a low level of phosphorylation on target proteins that regulate nerve activity. (This is because the ability of protein kinase A to phosphorylate these proteins is reduced, and because protein kinase A cannot phosphorylate DARPP-32 on threonine 34, which therefore cannot deactivate a phosphatase that also works on the target proteins.) Caffeine has some, although weak, effects in mice lacking DARPP-32. So Lindskog *et al.* suggest that this protein is not essential for, but rather intensifies and prolongs, the effects of low doses of caffeine.

Higher doses of caffeine would produce a long-lasting blockage of A_{2A} receptors, abolishing the adenosine-signalling pathway and in particular the activity of protein kinase A to the extent that no feedback loop would be required to inhibit this enzyme further. In other words, everything cooperates to abolish protein kinase A activity, leading to the lowest possible amounts of phosphorylated target proteins. So, behavioural responses to caffeine ultimately result from the activity of non-phosphorylated target proteins.

Of course, so far all of this work has been done in rodents. We don't yet know the

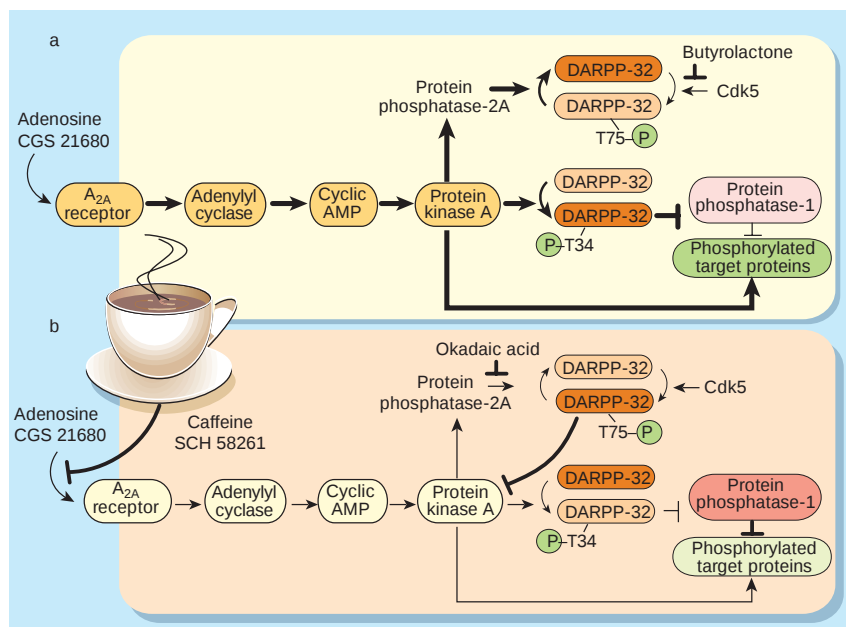


Figure 1 How coffee wakes up your nerve cells, based on work by Lindskog *et al.*¹. **a**, In the absence of caffeine. By activating A_{2A} receptors, adenosine increases adenylyl cyclase activity, leading to the formation of cyclic AMP and activation of protein kinase A. This enzyme in turn activates protein phosphatase-2A, leading to dephosphorylation of DARPP-32 at threonine 75 (T75). Protein kinase A also phosphorylates other target proteins. Finally, it phosphorylates DARPP-32 at T34, leading to the inhibition of protein phosphatase-1, which thus cannot dephosphorylate target proteins⁷. **b**, Partial blockade of A_{2A} receptors by low caffeine levels reduces the formation of cyclic AMP and the activation of protein kinase A, and hence of protein phosphatase-2A. The resulting increased phosphorylation of T75 converts DARPP-32 into an inhibitor of protein kinase A. This amplifies the effects of caffeine by a positive feedback loop. High caffeine levels virtually abolish the activation of protein kinase A. Further phosphorylation sites at serine amino acids add to the sophistication of DARPP-32 (not shown). CGS 21680, butyrolactone, okadaic acid and SCH 58261 are drugs used by Lindskog *et al.* to study these pathways. Thicker arrows and bars, and darker colours, indicate higher levels of activity.

equivalent doses of caffeine in human terms: the answer will await the ability to use positron emission tomography to study the occupancy of the A_{2A} receptor by caffeine in people's brains. But, from the information available so far, it seems likely that the concentrations of caffeine used by Lindskog *et al.*¹ are similar to those reached in our brains after a few cups of coffee. So the authors have led us to the heart of the matter: DARPP-32 keeps us going till the next coffee break by extending the effects of the last cup. Future efforts will identify the target proteins that are regulated by this signalling cascade. Together with insights gained from mice engineered to lack other intracellular signalling molecules, this knowledge should

provide an even better understanding of caffeine's effects. So, wake up, smell the fresh coffee and enjoy its effects for a long time, thanks to your dependable DARPP-32! ■

Jean-Marie Vaugeois is in the Unit of Experimental Neuropsychopharmacology, Faculty of Medicine and Pharmacy, Université de Rouen, 22 Boulevard Gambetta, 76183 Rouen Cedex 1, France.

e-mail: jean-marie.vaugeois@univ-rouen.fr

1. Lindskog, M. *et al.* *Nature* **418**, 774–778 (2002).
2. Fredholm, B. B., Battig, K., Holmen, J., Nehlig, A. & Zvartau, E. E. *Pharmacol. Rev.* **51**, 83–133 (1999).
3. Svenningsson, P., Nomikos, G. G. & Fredholm, B. B. *J. Neurosci.* **15**, 7612–7624 (1995).
4. Ledent, C. *et al.* *Nature* **388**, 674–678 (1997).
5. Chen, J.-F. *et al.* *J. Neurosci.* **21**, RC143 (1995).
6. Greengard, P. *Science* **294**, 1024–1030 (2001).
7. Svenningsson, P. *et al.* *Proc. Natl Acad. Sci. USA* **97**, 1856–1860 (2000).

Earth science

Breaking plates

J. Huw Davies

Seismic images suggest that oceanic plates in the northwest Pacific broke apart as they descended into Earth's mantle. That might explain the high magma output of some volcanoes in the region, and why others are extinct.

Some of the consequences of the motion of Earth's tectonic plates are quite evident: various mountain chains for instance. Others, such as events occurring deep in subduction zones where one plate plunges beneath another, can be seen only

indirectly through seismic imaging techniques. The approach allows the structure of Earth's interior to be reconstructed from observed seismograms, and it continues to pay dividends.

An example appears on page 763 of this

issue, where Levin *et al.*¹ argue that they have detected two instances of break-up of a subducting slab at the junction of two oceanic plates. There are indications^{2,3} that slab disintegration has occurred where oceanic and continental plates collide. But because earthquake activity is uniformly distributed with depth in subducting oceanic plates, it had seemed that purely oceanic plates maintain their integrity.

Levin *et al.* imaged the structure of the upper 200 km of Earth's mantle in the northwest Pacific, in the complex setting where the Aleutian and Kamchatka subduction zones meet (Fig. 1). South of the junction beneath Kamchatka, they identified a high-velocity feature co-located with earthquakes. This they interpret as cool, rigid subducting oceanic plate. North of the junction, velocities were slower than average, with no evidence of deep earthquakes, the implication being that there is no subducting slab there. This is an unexpected finding. In the past, subduction also occurred further north, the plate concerned being generated by ocean-floor spreading in the Komandorsky basin (Fig. 1a). If oceanic plates are rigid, the previously subducted plate north of the junction should still be attached to the Komandorsky basin, which has remained at the surface.

Levin and colleagues' conclusion is that the missing plate has broken off and descended deep into the mantle (Fig. 1b). Subduction zones are regions where Earth's surface returns to cool the hot interior. Counterintuitively, they are also the site of volcanism. The melting of rocks in Earth's interior that is required for this volcanism is thought to result from the entry of hydrous fluids into the hot overlying mantle from the descending oceanic crust^{4–6}. North of the junction lie volcanoes that are now extinct, and the inferred absence of oceanic crust there is taken by Levin *et al.* as an explanation: there is now no source of water to promote magmatism. The volcanoes have been extinct for at least five million years, so the plate must have become detached before then, taking oceanic crust away with it. If the plate continued its descent at the velocity of the Kamchatka slab (about 9 cm yr⁻¹), and assuming a dip of 45° or so, then in five million years it would have travelled 450 km and reached a depth of at least 300 km — beneath the region surveyed by Levin and colleagues.

Where oceanic and continental plates meet, the force that breaks a subducting slab is generated by the tendency of the buoyant continent to remain at the surface and that of dense oceanic plate to continue its descent. In the case studied by Levin *et al.*, if the fragment that was broken off was joined along its southern side to the subducting Pacific plate, then the driving force for detachment is clear: the much larger Pacific plate would have pulled the fragment down. Joining of the plates is not required, though, especially

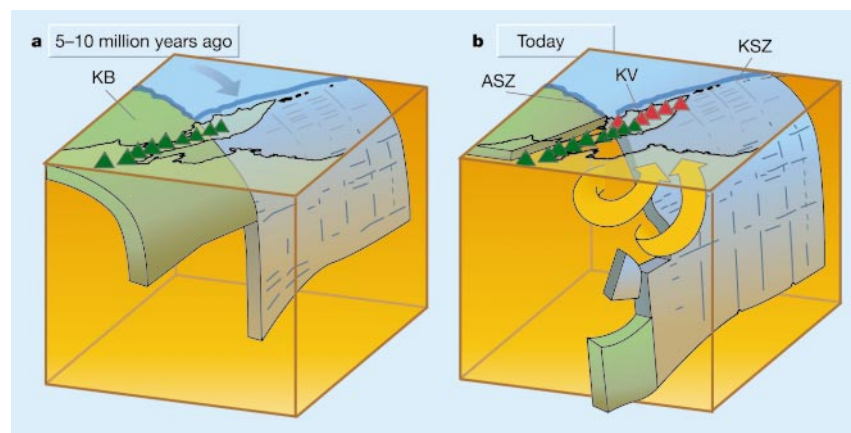


Figure 1 Tectonics in the northwest Pacific. **a**, A view from Siberia looking towards the Pacific, and interpretation of events 5–10 million years ago. The Pacific oceanic plate descends beneath south and central Kamchatka (blue), whereas plate produced by spreading in the Komandorsky basin descends beneath north Kamchatka (green). The green triangles show volcanoes then active. **b**, The Aleutian–Kamchatka junction today, interpreted from the work of Levin *et al.*¹. The plate beneath north Kamchatka seems to have broken off and fallen deep into the mantle. Volcanism usually occurs above subduction zones, the likely cause being water released from descended oceanic crust which enters the mantle and promotes melting. Levin *et al.* regard the lack of oceanic crust in the region above this detached segment as an explanation for the now-extinct overlying volcanoes (green triangles). The authors¹ also argue that a second segment of plate detached at depth beneath central Kamchatka, leading to stronger mantle upflow and the exceptional magma output of the Klyuchevskoy volcano. The red triangles indicate active volcanoes. KB, Komandorsky basin; KSZ, Kamchatka subduction zone; ASZ, Aleutian subduction zone; KV, Klyuchevskoy volcano.

as the two pieces of oceanic plate are of such different age. It might instead be that the Komandorsky basin plate is too buoyant to subduct, because it is young and hot, and that it pulled away from the subducted older plate. This process would have been enhanced if plates are weakest near the surface, at a region of maximum curvature that becomes cut by large faults just before subduction occurs⁷.

A further conclusion of Levin and colleagues is that a second piece of slab became detached at depth about two million years ago. The authors argue that the ensuing vigorous upflow of hot mantle, fed by water from the remaining ocean crust, explains the world-record magma output of the overlying Klyuchevskoy volcano (Fig. 1b). This is a plausible interpretation. But the mantle imaging is not good enough to be certain, and it is not the only explanation for the increased magmatic activity. Hot rock flowing around the side of the plate, which Levin *et al.* also predict is occurring, or increased hydrous-fluid input from subduction of part of the nearby Emperor seamount chain, are also possible causes. Other volcanoes with a high magma output also occur near junctions in subduction zones. Examples are Mount Fuji, where there is no evidence of underlying slab disintegration, and Mount Etna, where the activity is also attributed to enhanced upwelling flow⁸.

Tasks for the future are to identify the

forces driving ocean-slab disintegration and to image the broken pieces deep in the mantle beneath the northwest Pacific, and also to investigate similar places where Earth might have been negligent with its plates. Furthermore, the new work¹ has implications for researchers trying to model the dynamics and explain the geochemistry of Earth's mantle⁹. The break-up of subducting plates could introduce blobs¹⁰ — rather than sheets — of geochemically distinct material into the mantle mix that is being stirred by convection. Detailed observations such as those of Levin *et al.* should lead to more accurate estimates of both the geometry and the geochemical consequences of plate subduction — and so to better simulations of mantle convection and improved definition of the geochemical reservoirs. ■

J. Huw Davies is in the Department of Earth Sciences, Cardiff University, Cardiff CF10 3YE, UK. e-mail: huw@earth.cf.ac.uk

1. Levin, V., Shapiro, N., Park, J. & Ritzwoller, M. *Nature* **418**, 763–767 (2002).
2. Wortel, M. J. R. & Spakman, W. *Science* **290**, 1910–1917 (2000).
3. Davies, J. H. & von Blanckenburg, F. *Earth Planet. Sci. Lett.* **129**, 89–102 (1995).
4. Davies, J. H. & Stevenson, D. J. *J. Geophys. Res.* **97**, 2037–2070 (1992).
5. Gill, J. *Orogenic Andesites and Plate Tectonics* (Springer, Berlin, 1981).
6. Tatsumi, Y. *Geophys. Res. Lett.* **13**, 717–720 (1986).
7. Kanamori, H. *Phys. Earth Planet. Inter.* **4**, 289–300 (1971).
8. Gvirtzman, Z. & Nur, A. *Nature* **401**, 782–785 (1999).
9. Hofmann, A. W. *Nature* **385**, 219–229 (1997).
10. van Keken, P. E., Hauri, E. H. & Ballentine, C. J. *Annu. Rev. Earth Planet. Sci.* **30**, 493–525 (2002).

ventral cells, which normally express just Wg, to express Dpp as well⁴. Importantly, the additional legs generated in this way are composed mainly of wild-type cells, suggesting that the fates of these cells were changed by the Dpp- and Wg-expressing cells. Without this new source of Dpp and Wg, the wild-type cells would have contributed only to the normal legs.

Two models have been proposed to explain this result. The first is based on an analogy with Spemann and Mangold's classic demonstration⁵ of the existence of an 'organizer' — a group of cells that influence the behaviour of neighbouring cells — during amphibian development. According to this view, the intersection between the Wg and Dpp signals creates a distal 'organizer' in the leg disc⁴. It was proposed that this organizer is the source of a third signal that causes neighbouring cells to become distal; the farther a cell is from the organizer, the more 'proximal' its fate would be.

The alternative model posits that Wg and Dpp cause nearby cells to become distal directly, without inducing a third signal⁶. So, cells that receive the highest levels of both Dpp and Wg, such as those at the centre of the disc, would become distal, whereas cells receiving successively lower levels of both signals would take on successively more proximal fates.

Either model could account for the generation of bifurcated legs: a new source of Dpp and Wg would be expected to create a new proximal–distal axis whether indirectly (by creating an organizer) or directly. However, one limitation of the organizer model when it was originally suggested was that there was no candidate for the signal from the hypothesized organizer. By contrast, significant experimental support for the direct model came from showing that Dpp and Wg activate two genes along the proximal–distal axis directly, without inducing a third signal⁶. These genes, *Distal-less* (*Dll*) and *dachshund* (*dac*), are required for the generation of distal and intermediate regions, respectively³ (Fig. 1b). Thus, largely because of this work⁶, the prevailing theory has been that the combined and graded activities of Wg and Dpp directly produce the leg's proximal–distal axis.

But it takes more than *Dll* and *dac* to make a leg. After these genes have been turned on, several others, also required for distal leg fates, begin to be activated in subsets of the *Dll*-expressing leg region³. Although these genes cross-regulate each other⁷, it was not known whether they, too, would be targets for Wg and Dpp. Campbell⁴ and Galindo *et al.*² now show that Wg and Dpp are not required at the time when these genes are first turned on. Accordingly, whereas removing either Dpp or Wg function early in development truncates the legs, later removal has only minor effects on leg development².

Developmental biology

Signalling legacies

Richard S. Mann and Fernando Casares

Until now, the signals that control the development of the legs in insects and vertebrates have been thought to be different. But new work reveals similarities, which might have evolutionary implications.

We live in a three-dimensional world and, not surprisingly, the development of many animals — including ourselves — depends on the establishment of three body axes. Two of these, the head-to-tail (anterior–posterior) and back-to-front (dorsal–ventral) axes, are established in the egg before fertilization or, in some cases, as an immediate result of fertilization. In contrast, the proximal–distal axis, which extends from the base of each appendage (where it attaches to the body) to its tip, must be generated from scratch every time an animal grows a leg, arm or wing. How this axis is established has been the subject of debate for several decades, and work on leg and wing development in the fruitfly, *Drosophila melanogaster*, has provided many insights, at least for insects. Papers on page 781 of this issue¹ and in a recent issue of *Science*² take our understanding a step further,

by demonstrating an unexpected role for signalling through a protein called the epidermal-growth-factor receptor in fruitfly leg development.

Fly legs emerge from flat discs of tissue that are set aside during embryonic development³. At this stage two secreted signalling molecules, Wingless (Wg) and Decapentaplegic (Dpp), are needed to set up the proximal–distal axis³. These molecules are expressed as two opposing wedges in the disc, one ventral and the other dorsal (Fig. 1a, page 739). Where they meet, at the centre of the disc, will become the distal tip of the leg.

Not only are both signals required in the formation of the proximal–distal axis³, but additional axes can be generated experimentally by creating new sites at which Wg and Dpp intersect. For example, bifurcated legs (which have an additional proximal–distal axis) can be generated by forcing just a few

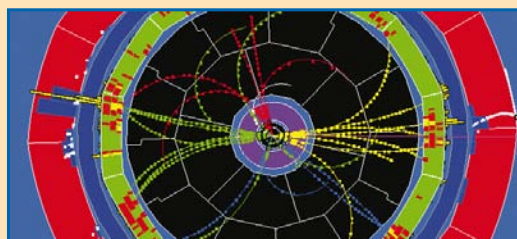
High-energy physics

Vote for Little Higgs

The Higgs boson has eluded detection for the moment, but theorists remain busy reconsidering what form or forms this object of much desire might take. The latest thinking was aired at the International Conference on High Energy Physics, held last month in Amsterdam.

The existence of the Higgs particle was first proposed in the 1960s. It is the manifestation of the Higgs mechanism that explains why fundamental particles, such as electrons and quarks, have mass. Mathematically, the particle can be formulated as part of the standard model of high-energy physics, or it could take a more exotic form that goes 'beyond the standard model' (BSM). But attempts to find evidence that the Higgs — any Higgs — exists have so far failed. An exhaustive search at CERN's Large Electron Positron (LEP) collider in Geneva ended in 2000 with the tantalizing hint of a signal, such as the candidate Higgs decay pictured here: the particle tracks shown in green and yellow may originate from a Higgs particle, created in an electron–positron collision at the centre. But no unequivocal signature has been identified.

So what now? Well, the experimentalists haven't given up. They will continue the search at the Tevatron collider, currently up and running at Fermilab, Chicago; and CERN's Large Hadron Collider (LHC) will take up the baton in 2007. Some theorists, however, have



gone back to the drawing-board. The basis of the Higgs mechanism remains, but tweaks to the model in the light of experimental data are altering what we might expect to see.

At the Amsterdam meeting, Martin Schmaltz (Boston Univ.) reported the results of a 'straw poll' he conducted among a gathering of theorists: what, he asked them, is the most exciting development in BSM theory? The clear winner was 'Little Higgs' theory, which many of the voters admitted to working on themselves. According to this theory, there should be one or more Higgs particles with masses around the 200-GeV range — the Higgs must be heavier than 114.4 GeV, according to the LEP data (<http://lepewwg.web.cern.ch/LEPEWWG>), and is probably lighter than a few hundred GeV. Moreover, the Little Higgs prediction is that there are other new particles to be discovered at about the 1,000-GeV scale. One of them could be a candidate particle for dark matter, the substantial mass that astronomical observations tell us is out there in the Universe, but is unseen.

So the goalposts have moved, but what is the significance of the new thinking? The theory of particles and interactions is beset with divergences — calculations that do not converge to a meaningful value. One solution is to invoke 'supersymmetry', a theory that predicts that every fundamental particle has a heavier, 'superparticle' partner. The Higgs mechanism and supersymmetry sit alongside each other quite happily, and physicists have even searched for a supersymmetric Higgs. But the Little Higgs theory can stand alone. It solves the problem of divergences in the standard model, without the need for supersymmetry.

A 200-GeV Higgs particle is an enticing prize, which Fermilab will no doubt race to find before CERN's LHC project comes into operation. The partner particles at 1,000 GeV are food for thought for proponents of the Linear Collider — a next-generation machine that has yet to get the go-ahead but that physicists hope will be operational within the next decade.

Alison Wright

If Wg and Dpp do not regulate the expression of these distal genes, what does? During vertebrate limb development, proteins from the fibroblast-growth-factor family are expressed at the distal edge of the growing limb. These proteins drive the formation of the proximal–distal axis by promoting the proliferation of underlying cells. They do this by activating receptors on the cell surface that are enzymes (receptor tyrosine kinases, RTKs), which transduce the signal through a well-characterized pathway to the cell nucleus⁸. Although fibroblast

growth factors do not seem to be involved in fly limb development, mutations in another RTK, the epidermal-growth-factor receptor (EGFR), cause the loss of the distal tip of the fly leg⁹. Campbell¹ has found that a gradient of EGFR signalling is likely to be at work, because a stepwise reduction in EGFR activity is correlated with a gradual loss both of leg structures and of leg-disc marker genes in a distal-to-proximal direction.

But which EGFR-activating proteins (ligands) are involved here? Both papers^{1,2} suggest the involvement of Vein, which is

expressed in cells at the centre of the disc, precisely where the Dpp and Wg signals intersect (Fig. 1c,d). This ligand also seems to be a target of Wg and Dpp but, once activated, no longer needs these signals for its expression². Yet removing Vein activity has only mild effects, if any, on the leg, so other ligands must also be involved. Indeed, Campbell shows that Rhomboid, which is required to process a different class of EGFR ligands, is also expressed in distal cells. In fact, it is not until Vein, Rhomboid and one of its relatives, Roughoid, are all simultaneously removed that a marked loss of distal structures occurs¹. So, although the specific ligands that Rhomboid and Roughoid activate are not known, the 'third signal' from the hypothesized distal organizer is likely to be a cocktail of EGFR ligands that are functionally redundant.

Thus, as often occurs in biology, both the direct and the organizer models seem to be correct, and each provides only part of the solution to how the proximal–distal axis is generated. The activity of Dpp and Wg apparently directly regulates *Dll* and *dac*, which define broad domains along the proximal–distal axis⁶. But the new papers^{1,2} provide strong evidence for a distal organizer, induced by Dpp and Wg, that is the source of EGFR ligands, which in turn works at a distance to help pattern the distal part of the leg.

The region of the leg that is patterned by this activity gradient — the tarsus — seems to be evolutionarily ancient¹⁰. This fact, together with the involvement of RTK pathways in both vertebrate and invertebrate^{1,2} leg development, makes it tempting to draw comparisons between the evolutionary origins of these limbs. However, the embryological differences are too plentiful to propose that the limbs are directly related to each other. Rather, the many developmental parallels, including RTK signalling, are more likely to reflect conserved regulatory relationships between signalling pathways that work together in many places in development. Alternatively, it is possible that vertebrate and invertebrate limbs were both derived from some type of outgrowth, perhaps not a limb¹¹, that grew from the body of a common ancestor. If this outgrowth had a primitive proximal–distal axis, it might have used many of the same genes and pathways that are active in the limbs we know today.

Beyond these evolutionary issues, other questions arise from this work. What are the EGFR ligands that Rhomboid and Roughoid process, and why are several different ligands required? Is activation of the EGFR pathway enough to generate an extra proximal–distal axis, or are other signals also required? To what extent is the regulation of the tarsal genes by the EGFR pathway direct and dependent on different thresholds of EGFR activity, as opposed to the result of cross-

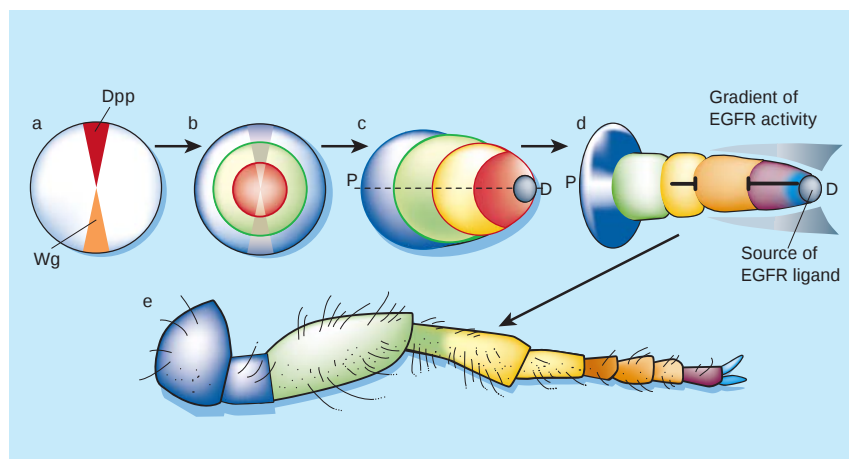


Figure 1 Taking steps to developing a leg. **a**, The Wingless (Wg) and Decapentaplegic (Dpp) proteins are expressed in wedges in the fly leg disc, which is set aside early in development. **b**, Wg and Dpp activate the *Dll* (red) and *dac* (green) genes at two different positions along the proximal–distal (P–D) axis. Proximally expressed gene-transcription factors are shown in dark blue. **c**, Growth of the leg disc creates a new domain in which *Dll* and *dac* expression overlaps (yellow). Vein and Rhomboid — which provide a source of ligands for the epidermal-growth-factor receptor (EGFR) — are activated by Dpp and Wg at the distal tip (grey spot). **d**, The source of EGFR ligands (grey spot) results in a gradient of EGFR activity (graded grey shading) that provides positional information for new domains of gene expression (purple and light blue). Cross-regulation by some of these factors (black bars) also contributes to the definition of these domains, which are shown as different shades of orange, purple and blue. Details are given in the new papers^{1,2}. **e**, A mature leg is patterned along its proximal–distal axis by a combination of the activity gradients of Dpp plus Wg, and EGFR.

regulation between these genes? We look forward to the next round of studies to provide an even more complete description of how to make a leg.

Richard S. Mann is in the Department of Biochemistry and Molecular Biophysics, Columbia University, 701 West 168th Street, New York, New York 10032, USA.

e-mail: rsm10@columbia.edu

Fernando Casares is at the Instituto de Biologia Molecular e Celular (IBMC), Rua do Campo Alegre 823, 4150-180 Porto, Portugal. e-mail: fcasares@ibmc.up.pt

1. Campbell, G. *Nature* **418**, 781–785 (2002); advance online publication, 24 July 2002 (doi:10.1038/nature00971).
2. Galindo, M. I., Bishop, S. A., Greig, S. & Couso, J. P. *Science* **297**, 256–259 (2002).
3. Morata, G. *Nature Rev. Mol. Cell Biol.* **2**, 89–97 (2001).
4. Campbell, G., Weaver, T. & Tomlinson, A. *Cell* **74**, 1113–1123 (1993).
5. Spemann, H. & Mangold, H. *Arch. Microsk. Anat. EntwMech.* **100**, 599–638 (1924).
6. Lecuit, T. & Cohen, S. M. *Nature* **388**, 139–145 (1997).
7. Kojima, T., Sato, M. & Saigo, K. *Development* **127**, 769–778 (2000).
8. Capdevila, J. & Izpisua Belmonte, J. C. *Annu. Rev. Cell. Dev. Biol.* **17**, 87–132 (2001).
9. Clifford, R. J. & Schupbach, T. *Genetics* **123**, 771–787 (1989).
10. Casares, F. & Mann, R. S. *Science* **293**, 1477–1480 (2001).
11. Shubin, N., Tabin, C. & Carroll, S. *Nature* **388**, 639–648 (1997).

Condensed matter

Scratching the Bose surface

Subir Sachdev

There are two distinct types of particles in nature: fermions and bosons. But it seems bosons may assume similar characteristics to fermion systems in the low-temperature regime typical of Bose–Einstein condensation.

Particles known as fermions (such as electrons) obey the Pauli exclusion principle, which decrees that only a single fermion can occupy a particular state, such as an orbital in an atom. But the second family of particles, bosons (such as helium atoms), have no corresponding restrictions on their occupation of states. This distinction in their individual properties also carries over to the collective properties of large numbers of fermions or bosons at low

temperatures. The typical quantum state of many fermions is a Fermi liquid, formed, for example, by the valence electrons in all metals: the electrons can move from atom to atom in this state and conduct electricity across macroscopic distances, albeit with a finite resistance. In contrast, bosons typically form a Bose–Einstein condensate, which is the basis of superfluidity in helium and its ability to flow perpetually with negligible dissipation.

Now, Paramekanti and colleagues¹, in an article to appear in *Physical Review B*, have proposed an innovative state for boson systems which shares many of its characteristics with the metallic state normally associated with fermions. According to Paramekanti *et al.*, strong interactions among the bosons entangle them in a delicate quantum-mechanical superposition that mimics the properties of fermions.

One of the defining properties of a Fermi liquid is its Fermi surface. In a metal, we can label the states available to the fermionic electron by its momentum vector **k**. The Fermi surface resides in **k**-space, and in the lowest-energy state of the metal, electrons occupy all states inside the Fermi surface, while those outside the surface remain unoccupied (Fig. 1a, overleaf). This surface is of particular physical importance because it defines the low-energy excitations observed experimentally: a fluctuation in the density of the electrons is created by moving an electron from an occupied state inside the Fermi surface to an unoccupied state outside the Fermi surface; probes that eject an electron from the metal (such as an energetic photon in a photoemission experiment) will act most efficiently on states just inside the Fermi surface. The Pauli exclusion principle, by enforcing single occupancy of fermionic states, is crucial in defining the concept of a Fermi surface, as electrons can only be removed (or added) from states inside (or outside) the Fermi surface.

The striking proposal of Paramekanti *et al.* is that bosons can also form a Bose liquid with a Bose surface, analogous to the Fermi liquid and Fermi surface. Clearly, as there is no Pauli exclusion principle for bosons, the Bose surface cannot be the boundary between occupied and unoccupied states. Instead, the Bose surface defines the locus of points in momentum space with low-energy excitations (Fig. 1b). Nevertheless, the excited states of the Bose liquid are defined by the Bose surface in a manner that is very reminiscent of the Fermi surface in a Fermi liquid: low-energy fluctuations in the density of the bosons involve rearrangement of the states near the Bose surface, as does an excitation that ejects a boson from the liquid.

Paramekanti *et al.* also discuss the conditions under which this fascinating Bose liquid may form. Weakly interacting bosons invariably collapse into the single state of a Bose–Einstein condensate, but merely turning up the strength of the repulsive interactions between the bosons is not enough to establish a Bose liquid. In fact, this leads to a state known as a Mott insulator, in which the bosons localize in a regular, crystalline arrangement. The transition between the Bose–Einstein condensate and the Mott insulator in a trapped gas of rubidium atoms was observed recently in a beautiful experiment by Greiner *et al.*²

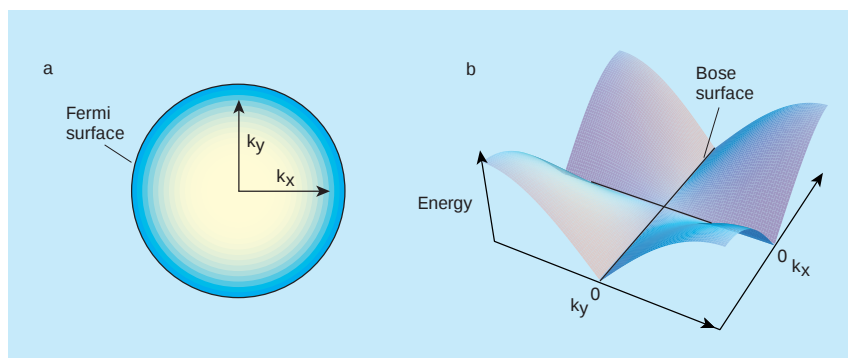


Figure 1 **Fermi and Bose surfaces in two dimensions.** a, Fermionic electrons in a metal obey Pauli's exclusion principle — there can be only one fermion in each quantum state. The Fermi surface marks the divide, defined by the particles' momentum components k_x and k_y , between occupied and unoccupied momentum states. b, Although bosons do not obey the exclusion principle, Paramekanti *et al.*¹ propose that they can form a Bose surface in a Bose liquid, analogous to the Fermi surface for fermions. In the spectrum of boson excitations in the Bose liquid, the excitation energy (vertical axis) vanishes near the Bose surface (which follows the lines $k_x = 0$ and $k_y = 0$): a Fermi liquid has a similar spectrum of excitations that vanishes on its circular Fermi surface.

To deter the bosons from forming a Bose–Einstein condensate or a Mott insulator, a more complex interaction between the bosons is necessary. In particular, a large contribution from ‘boson ring exchanges’ appears to be crucial. This feature arises because the quantum-mechanical description — or wavefunction — of the ground state contains a superposition of states in which pairs of bosons move in a correlated manner around a ring. In such a state, the positions of the bosons around the ring are uncertain, but a measurement that locates one boson at a particular site also specifies

the position of the second boson at another site on the ring.

Stimulated by the proposal of Paramekanti *et al.*, Sandvik *et al.*³ have already reported results from a computer simulation of the simplest quantum description of boson behaviour in two dimensions, including a large boson-ring-exchange term. So far, they have not found a state with a Bose surface, but have instead obtained a new form of Mott insulator in which the bosons crystallize on half the horizontal bonds of the lattice; this type of state had also been predicted⁴.

Paramekanti *et al.*¹ also discuss the prospects for experimental discovery of their Bose liquid state. They suggest that such a state might be formed by pairs of electrons (which act like composite bosons and are known as Cooper pairs) in the cuprate compounds that superconduct at high temperatures. Mason and Kapitlnik⁵ have recently reported an unexpected regime of metallic conduction in a disordered thin film of $\text{Mo}_{43}\text{Ge}_{57}$ in a magnetic field — the film is a superconductor in zero field, and a Bose liquid of Cooper pairs is an intriguing possibility for the metallic phase. And there are other competing theories^{6–8}, involving a more fundamental role for disorder in the film. It is clear that the resolution of the puzzle set by Paramekanti *et al.* brings the prospect of much exciting new physics.

Subir Sachdev is in the Department of Physics, Yale University, PO Box 208120, New Haven, Connecticut 06520-8120, USA.

e-mail: subir.sachdev@yale.edu

1. Paramekanti, A., Balents, L. & Fisher, M. P. A. Preprint cond-mat/0203171 (2002); <http://xxx.lanl.gov>. *Phys. Rev. B* (in the press).
2. Greiner, M., Mandel, O., Esslinger, T., Hansch, T. W. & Bloch, I. *Nature* **415**, 39–44 (2002).
3. Sandvik, A. W., Daul, S., Singh, R. R. P. & Scalapino, D. J. Preprint cond-mat/0205270 (2002); <http://xxx.lanl.gov>
4. Park, K. & Sachdev, S. *Phys. Rev. B* **65**, 220405 (2002).
5. Mason, N. & Kapitlnik, A. *Phys. Rev. B* **64**, 060504 (2001).
6. Spivak, B., Zyuzin, A. & Hruska, M. *Phys. Rev. B* **64**, 132502 (2001).
7. Dalidovich, D. & Phillips, P. *Phys. Rev. Lett.* **89**, 027001 (2002).
8. Das, D. & Doniach, S. *Phys. Rev. B* **64**, 134511 (2001).

Physiology

Muscle regulator goes the distance

Vertebrate skeletal muscles — the kind that enable a mouse to run on a wheel or you to dash for a bus — contain two types of fibre. Type I, or ‘slow-twitch’, fibres are fuelled mainly by oxidative metabolism, and can contract for sustained periods. Type II, ‘fast-twitch’ fibres generate rapid muscle contractions but are more prone to fatigue. On page 797 of this issue, Bruce Spiegelman and colleagues (*Nature* **418**, 797–801; 2002) describe how a protein that regulates gene expression can increase the type I fibre content, and consequently improve the stamina of isolated mouse muscles.

Previous studies have shown that the nuclear protein PGC-1 α is involved in controlling genes that participate in energy metabolism in mammalian cells. This suggested

that PGC-1 α might play a part in muscle physiology. Indeed, Spiegelman and colleagues have now found that if the protein is expressed in mouse skeletal muscles, more of the characteristically ‘redder’ type I muscle fibres are formed, with fewer type II fibres. The authors show that isolated PGC-1 α -expressing muscles contain proteins that are typical of type I fibres, and can sustain contraction for around twice as long as equivalent muscles from wild-type mice, in an experiment based on standardized electrically induced contraction.

These results tie in nicely with the fact that the expression of PGC-1 α can be induced naturally by exercise in experimental animals, and that exercise can convert type II fibres into type I fibres. Further work



will be needed to define the precise molecular signals that control PGC-1 α activity and to identify the genes regulated by this protein that lead to the formation of type I muscle fibres. It may be that, in the future, drugs that influence these

regulatory events will be useful in situations where muscle activity is impaired. Great care would have to be taken, however, as PGC-1 α seems to be involved in many other regulatory events in the body.

Richard Turner

Surprising strength of silkworm silk

Silk fibres produced by artificial reeling are superior to those that are spun naturally.

Commercial silkworm silk is presumed to be much weaker and less extensible than spider dragline silk, which has been hailed as a 'super-fibre'^{1–4}. But we show here that the mechanical properties of silkworm silks can approach those of spider dragline silk when reeled under controlled conditions. We suggest that silkworms might be able to produce threads that compare well with spider silk by changing their spinning habits, rather than by having their silk genes altered.

Typical commercial silkworm silk from *Bombyx mori* cocoons has a tensile strength of about 0.5 gigapascals (GPa), a breaking elongation of 15%, and a breaking energy (toughness) of $6 \times 10^4 \text{ J kg}^{-1}$ (ref. 5). *Nephila* spider dragline silk, on the other hand, typically has a strength of 1.3 GPa, a breaking elongation of 40%, and a toughness of $16 \times 10^4 \text{ J kg}^{-1}$ (ref. 4).

These mechanical measures vary considerably in both types of silk. For spider silk, this variability is due to spinning conditions, which are affected by the spider's body temperature and the speed of

drawing⁴. Silkworm silk is traditionally obtained from a natural cocoon that is spun by the moving silkworm, which accelerates and decelerates its head in arcs that are attached at points that correspond to each change of direction⁶.

We find that artificial reeling of silk from immobilized silkworms under steady and controlled conditions produces fibres that are superior to naturally spun ones. The silkworm, like the spider⁴, produces stronger, more brittle fibres at faster spinning speeds, whereas slower speeds lead to weaker, more extensible fibres (Fig. 1).

Force-drawn silkworm fibres compared favourably with *Nephila* spider dragline silk. For example, slow spinning at 4 mm s^{-1} gave a breaking elongation for *Bombyx* silk that is comparable to that of spider silk (37% and 35%, respectively). Faster spinning (13 mm s^{-1}) gave a breaking energy approaching that of *Nephila* silk ($12 \times 10^4 \text{ J kg}^{-1}$ for *Bombyx*; versus $16 \times 10^4 \text{ J kg}^{-1}$ for *Nephila*), although the breaking strength remained lower (0.7 GPa compared with 1.3 GPa). At even greater spinning speeds, silk toughness decreased, mainly because of loss of extensibility (Fig. 1).

Other differences between silkworm and spider silk may be linked to specific differences in the composition of the principal silk molecules², as well as to their arrangement during and after spinning¹. Depending on reeling speed, *Bombyx* silk is either strong or stretchable (Fig. 1), whereas spider dragline silk typically combines the two at the speeds that are characteristic of web construction⁴ ($10\text{--}20 \text{ mm s}^{-1}$).

Although spider silk fibres are studied in their native state⁴, silkworm silk must first be washed⁷ to remove most of the sericin gum coating the fibroin filaments from the 'bave' thread (Fig. 1, inset). Washing extraction is necessary for the thread to be unravelled from the cocoon (Fig. 2), and is generally assumed to weaken the thread⁸. Our experiments (each run four times) on silkworm silk filaments reeled under controlled conditions (12–24 filaments per treatment) showed that spinning conditions affect the silk's material properties more than washing.

For example, washing experimental fibres reeled at 4 mm s^{-1} changed neither the force required to break a thread (washed, $36 \pm 2 \text{ mN}$; native, $38 \pm 4 \text{ mN}$, $P = 0.34$, t -test) nor its breaking elongation (washed, $37 \pm 3\%$; native, $39 \pm 4\%$, $P = 0.53$). However, 'degumming' significantly increased viscoelastic recovery (washed, $40 \pm 3\%$; native, $29 \pm 2\%$, $P < 0.001$), as shown by the loading-unloading cycles of single fibres



Figure 2 A cocoon produced by the silkworm *Bombyx mori*. Changing the spinning conditions can improve the silk's quality.

stretched to 50% of typical breaking elongation values. Thus, more than anything else, washing increases the elasticity of fibroin filaments, presumably because of the combined action of sericin removal and water plasticization.

Our findings indicate that the mechanical properties of *Bombyx* silk, like those of spider silk⁴, depend crucially on spinning conditions — silkworm silks can be made stronger, stiffer and more extensible simply by adjusting the harvesting parameters. If we could reel silk straight from the silkworm, as from spiders, or if larvae could be bred to spin their cocoons faster and more evenly (with pretreatment washing appropriately modified), then the silkworm would produce fibres that might well give natural and artificially spun spider silks — genetically modified or not — a good run for their money.

Zhengzhong Shao*, Fritz Vollrath†

*Department of Macromolecular Science and Key Laboratory of Polymer Engineering of Education Ministry, Fudan University, Shanghai 200433, China

e-mail: zzshao@fudan.edu.cn

†Department of Zoology, Oxford University, South Parks Road, Oxford OX1 3PX, UK

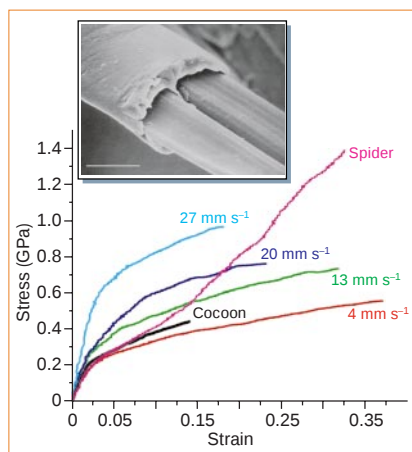


Figure 1 Comparison of silks drawn at different speeds from the silkworm *Bombyx mori*. Stress-strain curves of washed and degummed single-filament silkworm silk (motor-reeled at 25°C at the indicated speeds), *Nephila* spider dragline silk (20 mm s^{-1} at 25°C) and standard, degummed commercial silk from a cocoon spun by the animal in the natural 'figure of eight'⁹ at speeds oscillating between 4 and 15 mm s^{-1} at 20°C . The area under the stress-strain curve represents the energy that a fibre can take up before breaking, and thus indicates its toughness. Scale bar, $10 \mu\text{m}$. Immobilized silkworms ($n = 4$) were forcibly silked⁴, each providing 3–6 single filaments, which were tested in a stretching rig (force resolution, $30 \mu\text{N}$; time resolution, $< 5 \text{ ms}$ for 1 mN ; strain rate, 50% per min)⁴; further details are available from the authors. For silk 'degumming', a traditional aqueous solution standard wash of 1% sodium hydrogen carbonate¹⁰ was used, which led to a $30\text{--}40\%$ reduction in fibre diameter. Inset, unwashed native silkworm silk (photo courtesy of Giuliano Freddi).

1. Vollrath, F. & Knight, D. P. *Nature* **410**, 541–548 (2001).
2. Kaplan, D. L., Adams, W. W., Viney, C. & Farmer, B. L. *Silk Polymers: Materials Science and Biotechnology* (ACS, Washington, 1994).
3. Lazaris, A. et al. *Science* **295**, 472–476 (2002).
4. Vollrath, F., Madsen, B. & Shao, Z. *Proc. R. Soc. Lond. B* **268**, 2339–2346 (2001).
5. Wilding, M. A. & Hearle, J. W. S. in *Polymeric Materials Encyclopedia* Vol. 11 (ed. Salamone, J. C.) 8307–8322 (CRC, Boca Raton, Florida, 1996).
6. Miura, M., Pan, Z. J., Aoyama, S., Morikawa, H. & Mochizuki, S. *J. Seric. Sci. Jpn* **67**, 51–56 (1998).
7. Perez-Rigueiro, J., Elices, M. J., Llorca, J. & Viney, C. *J. Appl. Polymer Sci.* **82**, 1928–1935 (2001).
8. Perez-Rigueiro, J., Viney, C., Llorca, J. & Elices, M. *Polymer* **41**, 8433–8439 (2000).
9. Wiedbrauck, J. Z. *Tierpsychol.* **12**, 176–202 (1955).
10. Monti, P., Freddi, G., Bertoluzza, A., Kasai, N. & Tsukada, M. *J. Raman Spectrosc.* **29**, 297–304 (1998).

Competing financial interests: declared none.

Malaria

Thermoregulation in a parasite's life cycle

The life cycle of the malaria parasite *Plasmodium falciparum* goes through three developmental stages (schizogony, gametogony and sporogony), each of which presents different environmental constraints that must be met by an adaptive response in the parasite. Here we show that thermoregulation, in which the transcription of select RNAs is upregulated at cooler temperatures, is crucial to the developmental transition that occurs during the transmission of *P. falciparum* from human to mosquito. Our findings offer new insight into how the malaria parasite senses and reacts to its environment.

Switching between two sequence types of ribosomal RNA (designated A and S) occurs during the developmental cycle of *Plasmodium*^{1–5}. Although the resulting ribosomes still translate messenger RNA into protein, there are functional variations between the two types⁶. Polymorphisms in the sequences preceding two mature A-type rRNAs (A1, A2; GenBank accession numbers, AF503871 and AF5033868) and two mature S-type rRNAs (S1, S2; GenBank accession numbers, AF 503869 and AF503870) enabled us to follow the regulation of transcription of each precursor independently⁷ by real-time polymerase chain reaction. We found that precursors A1 and A2 are the predominant forms during schizogony, whereas S1 predominates during gametogony and S2 during sporogony.

We investigated the effect of temperature as a differential regulator of transcription. A single asexual parasite culture was split and the portions were incubated for 3 h at four different temperatures (Fig. 1a). The transcription of A1 and A2 remained relatively constant, but the rate of S-gene transcription was sensitive to temperature: at 42 °C, neither S gene was transcribed; S2 transcription increased 4.4-fold at 31 °C and 15-fold at 26 °C (monitored for comparison with S2 transcription at 37 °C). The S1 gene was affected by decreasing the temperature, but not as markedly.

We used nuclear run-on analysis⁸ to confirm this pattern of temperature regulation (Fig. 1b). We also measured the rates of transcription of three mRNAs, namely those encoding the *P. falciparum* proteins HRP III, MSP-1 and CSP (Fig. 1a). Our findings are consistent with the rRNA results in that the levels of the mRNAs associated with schizogony, HRP III and MSP-1, tend to fall off at reduced temperature, whereas CSP mRNA, which is associated with sporogony, increases. If the duration of culture is extended to 7 h, the difference

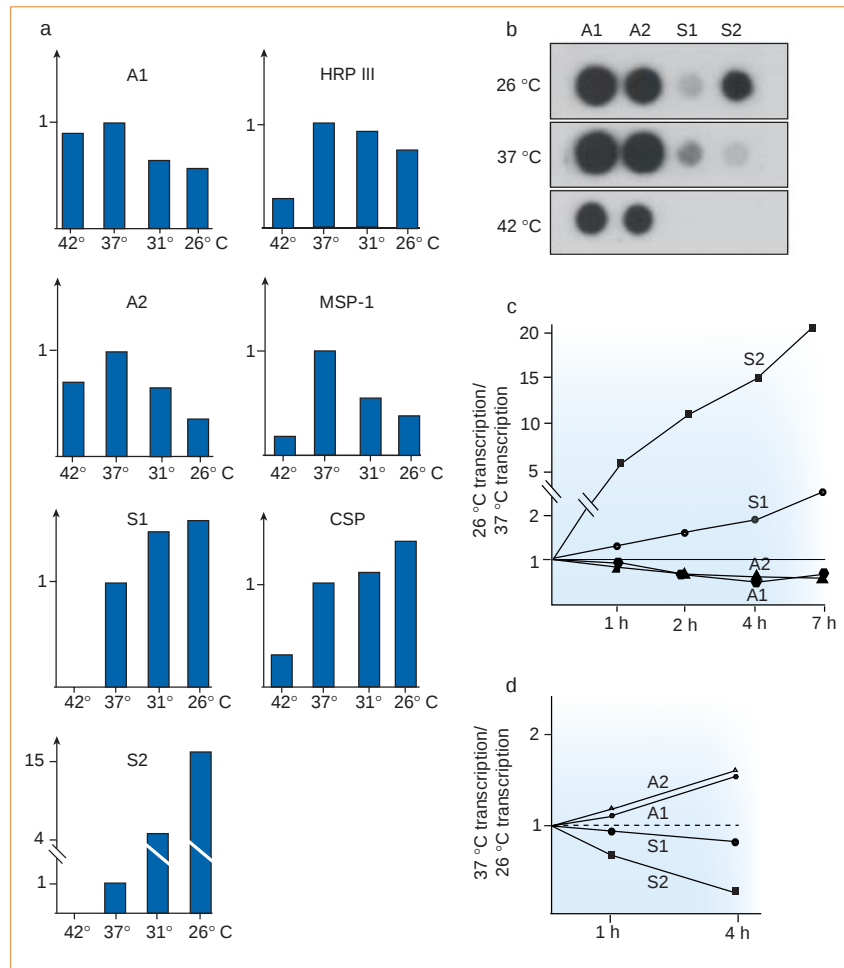


Figure 1 Effect of temperature on the transcription of selected genes in the malaria parasite. **a**, Transcription of individual genes (see text) at different temperatures likely to be encountered by *Plasmodium falciparum*. Results in asexual blood-stage parasites were determined by real-time polymerase chain reaction; the amount of transcription at 37 °C was taken as unity. **b**, Nuclear run-on analysis of the variation in transcription from ribosomal RNA genes with temperature. **c**, Changes over time in the relative levels of transcription at 26 °C and 37 °C of the different rRNA genes during the asexual blood stage. **d**, Changes over time in the relative levels of transcription at 37 °C and 26 °C of rRNA genes in mosquito-stage parasites.

in S2 transcription at 26 °C and at 37 °C is more than 20-fold (Fig. 1c).

This transcriptional control is evident by simply altering the temperature of the mosquito itself: we find that transcription of the S2 gene, but not of the other rRNA genes, is inhibited by warming (Fig. 1d).

A strong, cold-stimulated promoter might be responsible for this selective control of transcription. When we transfected *P. falciparum* with a chimaeric plasmid containing an S2 promoter region aligned with a fragment of the luciferase gene, we found that promoter control was fully operative over the range of temperatures at which *P. falciparum* is viable in culture (results not shown).

Temperature control is important with regard to malaria, given that vacillating ambient temperatures affect the rate of development of *P. falciparum* in the mosquito⁹. Our results provide a framework for understanding how the malaria

parasite responds to its environment, as well as indicating how this process might be investigated further.

Jun Fang, Thomas F. McCutchan

Growth and Development Section,
Laboratory of Parasitic Diseases,
National Institute of Allergy and Infectious
Diseases, National Institutes of Health, Bethesda,
Maryland 20892-0425, USA
e-mail: tmcutchan@niaid.nih.gov

- Waters, A. P., Syin, C. & McCutchan, T. F. *Nature* **333**, 74–76 (1988).
- Gunderson, J. H. et al. *Science* **238**, 933–937 (1987).
- McCutchan, T. F. et al. *Mol. Biochem. Parasitol.* **28**, 63–68 (1988).
- Li, J., Wirtz, R. A., McConkey, G. A., Sattabongkot, J. & McCutchan, T. F. *Mol. Biochem. Parasitol.* **65**, 283–289 (1994).
- Rogers, M. J. et al. *RNA* **2**, 134–145 (1996).
- Velichutina, I. V., Rogers, M. J., McCutchan, T. F. & Lieberman, S. W. *RNA* **4**, 594–602 (1998).
- Waters, A. P. et al. *J. Biol. Chem.* **272**, 3583–3589 (1997).
- Rosalina, M. L. et al. *Mol. Biochem. Parasitol.* **99**, 193–205 (1999).
- Noden, B. H., Kent, M. D. & Beier, J. C. *Parasitology* **111**, 539–545 (1995).

Competing financial interests: declared none.

A physical map of the mouse genome

Simon G. Gregory*, Mandeep Sekhon†, Jacqueline Schein‡, Shaying Zhao§, Kazutoyo Osoegawa||, Carol E. Scott*, Richard S. Evans*, Paul W. Burridge*, Tony V. Cox*, Christopher A. Fox*, Richard D. Hutton*, Ian R. Mullenger*, Kimbly J. Phillips*, James Smith*, Jim Stalker*, Glen J. Threadgold*, Ewan Birney¶, Kristine Wylie†, Asif Chinwalla†, John Wallis†, LaDeana Hillier†, Jason Carter†, Tony Gaige†, Sara Jaeger†, Colin Kremitzki†, Dan Layman†, Jason Maas†, Rebecca McGrane†, Kelly Mead†, Rebecca Walker†, Steven Jones‡, Michael Smith‡, Jennifer Asano‡, Ian Bosdet‡, Susanna Chan‡, Suganthi Chittaranjan‡, Readman Chiu‡, Chris Fjell‡, Dan Fuhrmann#, Noreen Girn‡, Catharine Gray‡, Ran Guin‡, Letticia Hsiao‡, Martin Krzywinski‡, Reta Kutsche‡, Soo Sen Lee‡, Carrie Mathewson‡, Candice McLeavy‡, Steve Messervier‡, Steven Ness‡, Pawan Pandoh‡, Anna-Liisa Prabhu‡, Parvaneh Saeedi‡, Duane Smailus‡, Lorraine Spence‡, Jeff Stott‡, Sheryl Taylor‡, Wesley Terpstra‡, Miranda Tsai‡, Jill Vardy‡, Natasja Wye‡, George Yang‡, Sofiya Shatsman§, Bola Ayodeji§, Keita Geer§, Getahun Tsegaye§, Alla Shvartsbeyn§, Elizabeth Gebregeorgis§, Margaret Krol§, Daniel Russell§, Larry Overton§, Joel A. Malek§, Mike Holmes§, Michael Heaney§, Jyoti Shetty§, Tamara Feldblyum§, William C. Nierman§, Joseph J. Catanese||, Tim Hubbard*, Robert H. Waterston†, Jane Rogers*, Pieter J. de Jong||, Claire M. Fraser§, Marco Marra‡, John D. McPherson† & David R. Bentley*

* The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

† Genome Sequencing Center, Washington University School of Medicine, St Louis, Missouri 63108, USA

‡ Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia V5Z 4E6, Canada

§ The Institute for Genomic Research, Rockville, Maryland 20850, USA

|| Children's Hospital Oakland Research Institute, Oakland, California 94609, USA

¶ EMBL—European Bioinformatics Institute, Hinxton, Cambridge CB10 1SD, UK

Department of Electrical Engineering, Washington University, St Louis, Missouri 63130, USA

A physical map of a genome is an essential guide for navigation, allowing the location of any gene or other landmark in the chromosomal DNA. We have constructed a physical map of the mouse genome that contains 296 contigs of overlapping bacterial clones and 16,992 unique markers. The mouse contigs were aligned to the human genome sequence on the basis of 51,486 homology matches, thus enabling use of the conserved synteny (correspondence between chromosome blocks) of the two genomes to accelerate construction of the mouse map. The map provides a framework for assembly of whole-genome shotgun sequence data, and a tile path of clones for generation of the reference sequence. Definition of the human–mouse alignment at this level of resolution enables identification of a mouse clone that corresponds to almost any position in the human genome. The human sequence may be used to facilitate construction of other mammalian genome maps using the same strategy.

The recent revolution in large-scale analysis of genomes has already yielded near-complete DNA sequences of a diverse range of organisms including bacteria, a yeast, a worm, a fly and humans^{1–7}. The sequence of the mouse genome will have a huge impact on biological research and human health. It will provide critical information and reagents for use in mouse experimental models. It will become possible to unravel the mechanisms of complex mammalian biological processes and human disease. The mouse genome sequence will provide the first opportunity to compare the complete sequence and organization of the human genome with that of another mammal. It will aid discovery and annotation of gene structures and other functionally important sequences in both genomes.

Assembly of each large genome sequence to date has been underpinned by production of a comprehensive map of overlapping large-insert bacterial clones (for example, cosmid⁸ or bacterial artificial chromosome (BAC) clones⁹)^{10–13} for sequencing, and, in some cases, for integration with whole-genome shotgun sequence data (refs 5, 7; see also review in ref. 14). The clones provide substrates for finishing each section of the genome sequence separately, making it easier to eliminate errors and to achieve a final accuracy of more than 99.99% (refs 14–16). The study of other large genomes requires physical maps of a similar standard for sequencing and biological studies.

Comparisons between genomes reveal homologous sequences, for example within individual genes, that reflect their common evolutionary origin and subsequent conservation. Similarity between genomes is also evident at the level of long-range sequence organization. A 'conserved segment' (also called a 'conserved linkage') refers to a region where the order of multiple genes on a

single chromosome segment (and hence the linkage between them) is the same in both species. A 'conserved synteny' refers to a region where the chromosomal location of multiple genes is conserved, but not necessarily their precise order^{17–19}. In general, the degree of similarity at all levels is higher between species that diverged more recently from a common ancestor. The long-range organization of the mouse and human genomes is very similar, and may comprise about 180 conserved syntenic regions^{6,17}. Detailed studies in some smaller regions suggest that precise gene order has been conserved, confirming the existence of conserved segments^{20,21}. The similarity in sequence organization between these two genomes suggests that the human genome can be used as a framework to map the mouse genome^{20,21}, and possibly other mammalian genomes as well. Alignment of the different genome maps over their entire length would simultaneously define the syntenic relationship between them at a new level of resolution, and accelerate the process of constructing the mouse clone map.

Using the human sequence as a framework

We constructed a physical clone map of the mouse genome in two phases. In the first phase, we compared restriction digest patterns ('fingerprints') of 305,716 BAC clones²². We identified overlaps between clones on the basis of similarity between fingerprints and used this information to construct 7,587 contigs of overlapping clones. In parallel, we determined sequences at the ends of the mouse DNA fragments cloned in the BACs (that is, BAC end sequences, or BESs). We aligned the mouse BAC contigs to the human genome sequence by identifying matches between human sequence and mouse BESs. We extended and joined contigs where possible after re-examining the fingerprint data. Overlapping fin-

gerprints were required to support each join, while juxtaposition of the mouse clone contigs along the human genome greatly accelerated the manual editing process. This phase resulted in generation of a human–mouse homology clone map. In the second phase, we used a set of independently mapped mouse markers (available in existing genetic and radiation hybrid maps of the mouse) to position the BAC contigs in the mouse genome. After further manual contig editing was carried out, we generated a mouse clone map comprising 296 contigs.

An illustration of the human–mouse homology clone map produced in the first phase is shown in Fig. 1. Within a 1.6-megabase (Mb) region of human chromosome 6 (6q16.1), 11 of the 15 segments of human sequence match to 29 of the BESs within a mouse BAC contig. This contig is located on mouse chromosome 4 (Mmu4). The syntenic relationship between the two genomes in this area was previously indicated by the proximity of the two loci *CNR1* and *GABRR1* (human) and their corresponding homologues *Cnr1* and *Gabrr1* (mouse) (also shown in Fig. 1). In our analysis here, a further 29 homologous crosslinks between these two loci are added.

We found 51,486 homologous crosslinks between the two genomes during this analysis, which corresponds to one match per 54 kb on average (assuming a 2.8-Gb mouse genome; see Methods). These matches were derived from a complete set of 453,962 BESs. Some 73% of the BESs contained more than 100 base pairs (bp) of contiguous unique sequence; we excluded the rest from the analysis.

In all, 1,326 of the BESs matched human chromosome 20 sequence. We found that 19% of them matched putative coding regions (163 matches to known genes, 80 to new genes, 8 to putative genes and 1 to a pseudogene). The rest were matches to other conserved sequences in introns (32%) and intergenic regions (49%). Human–mouse sequence homologies in non-coding regions have also been observed in other studies^{20,23,24}. From this analysis it is clear that non-coding homologies contributed significantly to the alignment.

Of the clones in the human genome tile path, 88% are collinear with the mouse BAC map (Table 1). For individual human chromosomes, coverage by aligned mouse contigs exceeds 80% on all except chromosome 19 (61%) and the Y chromosome (0%). Lower coverage of these human chromosomes can be accounted for by several factors. First, we used only BES matches with unique locations in the human tile path for the alignment. BESs matching regions of human sequence containing dispersed low-copy repeats were filtered out of the analysis. Such regions are known on human chromosome 19 (notably including zinc-finger-containing genes among others^{21,25}). Second, a significant fraction of chromosome 19 clones (49 in all) are included in the total in Table 1 but contain pericentromeric repetitive sequence, and any BES matches to them were excluded. A further 118 clones on chromosome 19 were covered by contigs aligned in our study, but we did not score them in Table 1 because they contained no high-scoring matches to BESs. This is reflected in the low overall density of BES matches on

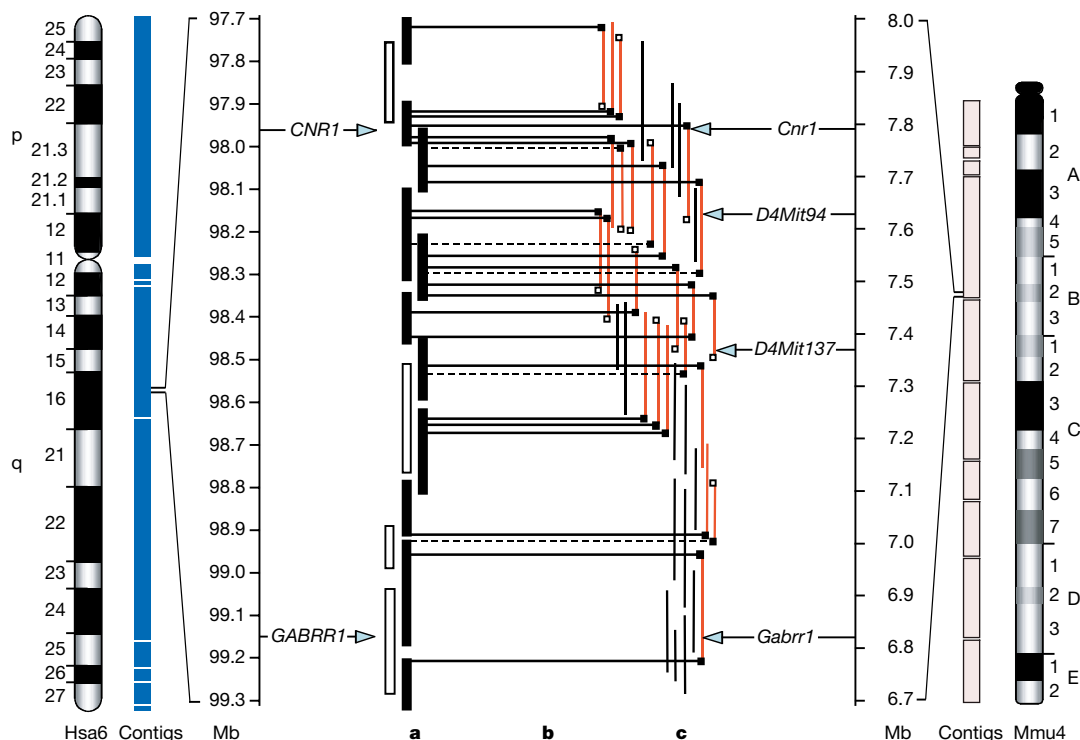


Figure 1 Construction of human–mouse homology clone map. Alignment between part of human chromosome 6 (Hsa6) and mouse chromosome 4 (Mmu4). Clone contigs are shown as blue and pale pink boxes, respectively. A 1.6-Mb interval is enlarged, showing part of Hsa6q16.1 aligned to a 1.3-Mb mouse BAC contig. The accompanying megabase scale for Hsa6 starts from one end of the chromosome (ptel, 0 Mb; qtel, 172 Mb), whereas the megabase scale shown for the mouse contig starts at one end of that contig. **a**, Human accessions. Filled and clear boxes denote accessions with and without matches to BESs, respectively. **b**, Matches between human sequence and BESs. Black lines denote high-homology BLAST matches and dashed lines indicate medium-homology BLAST matches (see Methods for details). **c**, Mouse BAC clone contig. All clones having at least one BES

that matches a human accession are shown in red (black and white squares at the ends of clones denote BESs with and without matches to human accessions, respectively). Black lines represent a subset of the other clones in the contig (the rest have been omitted for clarity, but can be viewed in Supplementary Information). Previous alignment of the two human loci *CNR1* and *GABRR1* with their mouse homologues can be viewed in the NCBI versus MGD alignment (<http://www.ncbi.nlm.nih.gov/Homology>). Another human locus (*RNGTT*) lies between *CNR1* and *GABRR1* in the human sequence, and its murine homologue (*Rngtt*) is present in the NCBI versus MGD alignment; however, *Rngtt* has not been placed in the mouse clone map yet. *D4Mit94* and *D4Mit137* are additional mouse markers that have been placed in the clone map by this study.

Table 1 Summary statistics of human–mouse homology clone map

Human chromosome number	Human sequence available (Mb)	Human–mouse sequence matches		Number of clones in human tile path	Clones in human tile path that are covered in mouse	
		Total	Per Mb		Number	Per cent
1	239	4,888	20.5	2,179	2,060	95
2	242	4,783	19.8	1,951	1,675	86
3	204	3,667	18.0	1,700	1,493	88
4	189	2,747	14.5	1,612	1,510	94
5	182	2,892	15.9	1,973	1,682	85
6	176	3,264	18.5	1,800	1,666	93
7	161	2,957	18.4	1,518	1,342	88
8	143	2,310	16.2	1,210	1,040	86
9	114	2,505	22.0	963	883	92
10	140	2,531	18.1	1,132	1,014	90
11	139	2,606	18.7	1,134	1,013	89
12	137	2,206	16.1	1,022	943	92
13	99	1,318	13.3	851	798	94
14	90	2,103	23.4	655	642	98
15	82	1,721	21.0	671	573	85
16	81	1,470	18.1	726	666	92
17	81	1,585	19.6	656	612	93
18	79	1,226	15.5	616	564	92
19	46	507	11.0	841	512	61
20	61	1,326	21.8	632	608	96
21	34	507	14.9	475	432	91
22	34	439	12.7	345*	279	81
X	149	1,925	12.9	1,555	1,276	82
Y	26	0	0.0	208	0	0
All	2,928	51,483	17.5	26,425	23,283	88

*For chromosome 22, the finished sequence was divided into abutting 100-kb segments for this analysis.

this chromosome (Table 1) and may be the result of chromosome-specific low-copy repeats, as observed previously^{21,25}. We were not able to align any mouse contigs to the human Y chromosome. Two reasons may account for this. First, much of the human Y sequence is repetitive (with Y-specific, X–Y and autosome–Y homologies), and mouse BES matches to these regions would have been excluded. Second, only 39% of the BESs were derived from a male mouse library, and we estimate that only about 0.4% of the BESs were therefore derived from the mouse Y chromosome.

Of the total coverage of the mouse BAC map (in 211 contigs), 97% (2,658 Mb) is aligned to the human genome sequence. The

other 3% (84 Mb, in 85 contigs) may not have been aligned at this stage for several reasons. First, the BESs may not find matches in the human sequence because the human sequence is incomplete. In a recent analysis, 3% of a curated set of 10,212 full-length human complementary DNAs (taken from the set in the ‘reference sequence project’, <http://www.ncbi.nlm.nih.gov>) were found to have no match in the human sequence (using BLAT²⁶ to identify matches with more than 95% identity between cDNA and the human genome sequence (NCBI Build 28); J.Kent and D. Haussler, personal communication). This fraction is sufficient to account for all of the unplaced mouse contigs. Second, most of the unplaced contigs are

Table 2 Summary statistics of mouse physical clone map by chromosome

Mouse chromosome number	Size estimate (Mb)		Total contigs			
	Old*	New	Number	Coverage (Mb)	Percentage of old size estimate	Percentage of new size estimate
1	216	196	10	211	98	108
2	209	190	13	186	89	98
3	180	163	8	173	96	106
4	177	160	10	145	82	90
5	170	154	24	156	92	101
6	166	151	16	157	95	104
7	156	141	14	143	92	101
8	149	135	12	140	94	104
9	144	131	15	127	88	97
10	145	131	9	135	93	103
11	142	129	6	107	75	83
12	146	132	8	126	86	95
13	131	119	11	132	101	111
14	134	121	8	130	97	107
15	122	111	7	108	89	98
16	114	103	7	101	88	97
17	116	105	6	95	82	91
18	116	105	6	96	83	91
19	82	74	4	65	80	88
X	187	170	30	132	71	78
Y	86	78	4	14	17	18
Subtotal			228	2,679	87	96
Unlocalized			68	60		
Total	3,088	2,800	296	2,739	89	98

*Taken from ref. 47.

short and contain few BESs. The production of additional mouse sequence in these contigs may allow identification of homology crosslinks, and hence alignment of the contigs to the corresponding region in the human genome. Third, there may be insufficient homology between the two genomes in some regions.

The mouse clone map

The mouse genome comprises 19 autosomes plus the two sex chromosomes, X and Y. In contrast to the human genome, all the mouse chromosomes are acrocentric (that is, the centromere is at or near one end). To construct the clone map for each mouse chromosome, we made use of the available genetic map (<http://www-genome.wi.mit.edu/cgi-bin/mouse/index>) of 6,559 gene-specific or simple sequence length polymorphism (SSLP) markers^{27–29}, and also a radiation hybrid (RH) map³⁰ of 13,558 markers, which included a subset of 2,446 of the markers previously ordered in the genetic map. This provided a total unique set of 18,379 markers. We placed 48% (8,789) of these markers within the mouse BAC contigs using the BESs or matches to available mouse genomic sequence, or by hybridization of markers to arrayed BAC clones^{31–33}. We then added another 8,203 markers to the map by hybridization to the BACs. As a result, we were able to assign a unique position and orientation on a mouse chromosome to 203 contigs (containing 2,631 Mb, or 96% of the map coverage), and a chromosome assignment (but not a position) to 25 more contigs (containing a total of 47 Mb). A further 68 contigs (containing 60 Mb) have no chromosomal assignment. A summary of the mouse map by chromosome is shown in Table 2.

Most mouse BAC contigs contained multiple mouse markers (average 57 markers per contig). We compared the order of markers within the contigs to the order suggested previously by the genetic and RH data. Overall, there was excellent agreement between the order of markers in each map, as illustrated for example on mouse chromosome Mmu2 (see Fig. 2; and Supplementary Information for all chromosomes). There was also concordance of the marker order in the clone map with that of the European Union RH map

(http://www.genoscope.cns.fr/externe/English/Projets/Projet_ZZZ/rhmap.html). The availability of integrated map information permitted assessment of the remaining 27 minor conflicts. We found two conflicts between the clone map and both the genetic and RH maps. The other 25 were conflicts solely with the RH map, and occur in regions where the genetic map offers limited or no resolution. For example, we inspected the RH data in the 80–90-Mb region of Mmu2 (Fig. 2) and found that the number of obligate breaks between the framework markers can be reduced if these markers are arranged according to the order in the clone map. We made the same observation in all eight cases examined in detail (see Supplementary Information). In general, it appears that the conflicts are caused by local inversions owing to a few (typically one or two) suboptimally located framework markers in the RH map. This in turn causes the order of all other markers previously placed in the RH map relative to that framework marker to be in conflict with their order in the clone map.

We carried out an additional check of contig assembly by comparing ‘virtual’ fingerprints calculated from finished mouse

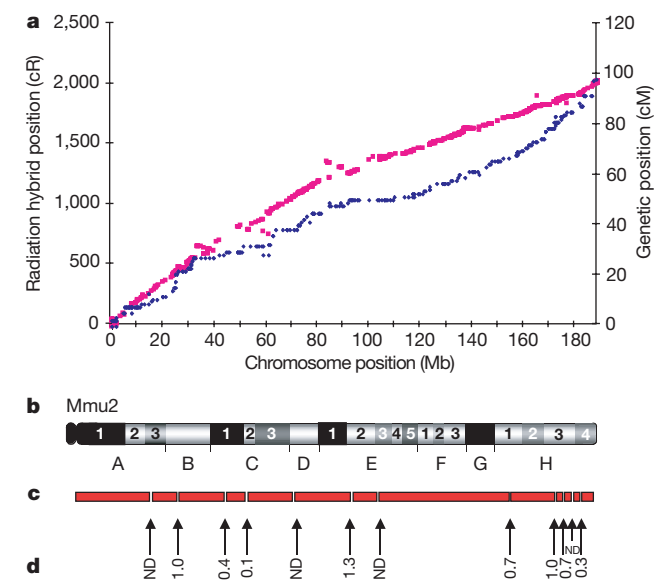


Figure 2 Comparison of radiation hybrid and genetic maps to the physical map of mouse chromosome 2. **a**, Radiation hybrid (pink) and genetic (dark blue) markers are plotted against their respective positions in the physical map of chromosome 2. **b**, Mouse chromosome 2 ideogram. **c**, **d**, Mouse BAC contigs (**c**) and gap sizes (**d**), estimated from bridging human sequence as described in the text. cR, centiRays; cM, centimorgans; ND, not determined.

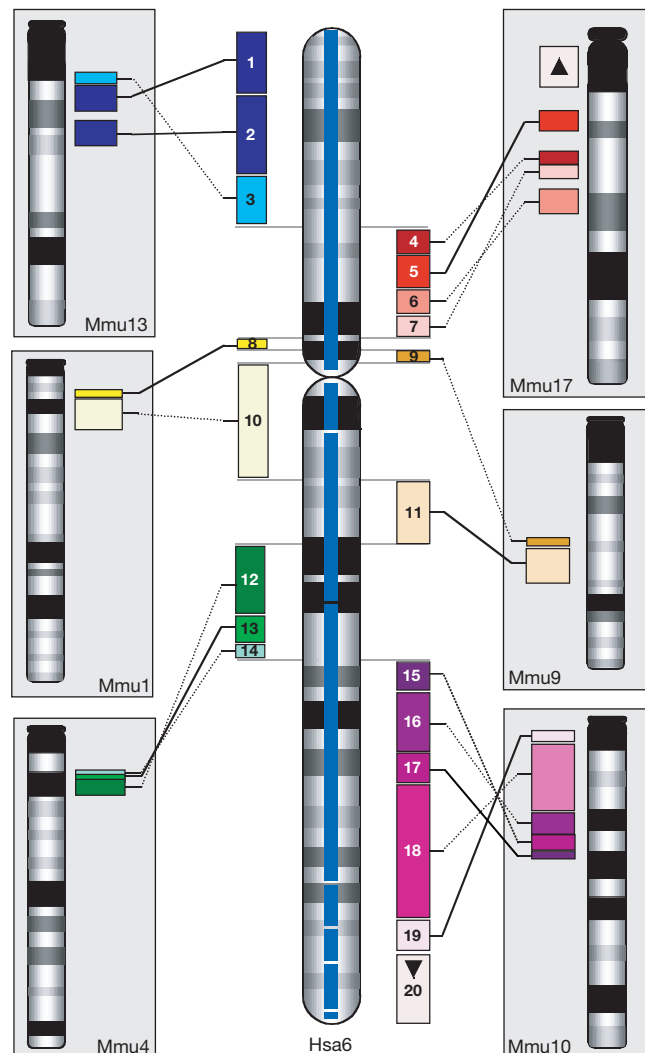


Figure 3 Conserved segments between human chromosome 6 (Hsa6) and the mouse genome. Blue bars in the centre of the chromosome ideogram represent the current extent of contig coverage in the human map. Numbered accession blocks correspond to Table 3 and denote conserved synteny with the mouse chromosomes. Dotted lines joining coloured blocks between the two species indicate conserved syntenic blocks inverted with respect to each other relative to the orientation of each chromosome as drawn.

Table 3 Conserved segments between human chromosome 6 and mouse

Segment	Human ptel–qtel number*	GenBank accession numbers	Human segment size (Mb)	Human–mouse homology hits	Hits per Mb	Mouse chromosome number	Mouse segment size (Mb)
6.1	3–97	AL365272–AL031785	9.0	220	24	13	9.9
6.2	99–207	AL031904–AL022726	10.3	303	29	13	9.0
6.3	209–302	AL008627–AL049543	9.5	146	15	13	9.9
6.4	310–374	AL022727–Z97183	5.6	89	16	17	3.4
6.5	376–432	AL161903–AL035690	5.1	116	23	17	4.4
6.6	434–453	AL136087–AL031778	1.8	37	21	17	1.9
6.7	458–537	AL139331–AL390247	6.5	178	27	17	6.6
6.8	560–582	AL360175–AL109918	2.7	70	26	1	3.1
6.9	582–613	AL109918–AL589796	3.1	57	18	9	2.9
6.10	620–771	AL031779–AL365232	12.4	204	16	1	13.2
6.11	773–900	AL603910–AL136082	11.1	257	23	9	10.9
6.12	918–998	AL096817–AL513186	7.0	174	25	4	7.5
6.13	1,005–1,031	AL607077–AL589740	2.9	45	16	4	2.7
6.14	1,033–1,044	Z84482–AL390959	1.1	24	22	4	1.7
6.15	1,048–1,120	AL080285–AL645528	6.8	120	18	10	8.3
6.16	1,121–1,227	AL355586–AL121953	9.2	157	17	10	10.4
6.17	1,229–1,297	AL354716–AL512283	6.2	141	23	10	7.0
6.18	1,301–1,579	AL591428–AL078581	26.1	616	24	10	27.7
6.19	1,584–1,637	AL355497–AL591419	3.9	65	17	10	4.1
6.20	1,639–1,797	AL591499–AL031259	15.0	246	16	17	11.4
Total			155.3	3,265	21		156.0

*Numbered arbitrarily from 1 (nearest the telomere of the short arm) to 1,800 (nearest the long-arm telomere).

sequence with the experimental fingerprints in the database. This was possible for 762 clones, each with more than 100 kb of finished sequence available. All virtual fingerprints were matched with high confidence to an experimentally derived fingerprint in the database, thus confirming the accuracy of the fingerprint data. Some 97% of them also matched to the expected position in contigs, confirming the assembled BAC map. The remaining 3% (20 clones) matched to other unique locations. These can be attributed to errors in tracking clone names during the project, and are generally resolved later by checking the overlaps between clones using sequence information as it becomes available.

Assuming the mouse genome is about 2.8 Gb in total (see Methods), coverage of the mouse genome in mapped BACs is virtually complete: 296 contigs of average size 9.3 Mb cover an estimated 2,739 Mb. The estimated coverage of each mouse chromosome by BAC contigs is listed in Table 2. Four contigs were placed on the mouse Y chromosome. The low coverage is partly due to the lack of previously mapped markers, and partly because only part of the fingerprinted BAC collection was derived from a male mouse. A number of the unplaced contigs contain clones derived solely from the male mouse BAC library, and may also map to the mouse Y chromosome when more data is available. The map continues to be updated, and the current version can be viewed at <http://www.ncbi.nlm.nih.gov/genome/guide/mouse> (choose 'MapView') or at http://www.ensembl.org/Mus_musculus/cytoview (use the 'Jump to chr.' box to select chromosome; if searching specific features, select 'cytoview' option to view the map). The archive of the map at the time of publication is available as Supplementary Information. The level of continuity of the mouse clone map is similar to that of the current human genome map, which comprises 279 contigs covering approximately 2,928 Mb (as of March 2002; International Human Genome Sequencing Consortium, I. Vastrik, personal communication).

There are 275 gaps in the mouse clone map. Many of these gaps occur where there is a break in synteny between the two genomes. No direct measurement of the existing gaps has yet been attempted (for example by fluorescent *in situ* hybridization to chromosomes or DNA fibres). However, 83 gaps in the mouse map are bridged by human contigs. The total length of human sequence across these gaps is 51.9 Mb. If we assume collinearity between the mouse and human genomes across these gaps, then the average size of each gap in the mouse map is 0.63 Mb. On mouse chromosome 2, for example, 9 of the 12 gaps between contigs are bridged by segments of the human sequence ranging from 0.3 to 1.3 Mb (see Fig. 2). More

practically, the tracts of human sequence that bridge gaps in the mouse map may provide a source of probes (for example, selected from annotated exons in the human sequence) to screen for new mouse clones and thus assist in closing the gaps. Conversely, mouse clones that are collinear with gaps in the human tile path may be used to assist closure of the human map. Our study indicates that 132 of the 244 gaps in the human map (March 2002 release) are bridged by mouse contigs and might be closed in this way.

Human–mouse homology maps

Alignment of the human and mouse clone maps using 51,486 homology crosslinks allowed us to define the segments that are conserved in both genomes at a new level of resolution. In many cases we were able to locate the boundary between adjacent segments to within one or a few clones. For example, the sequence of human chromosome 6 contains 1,800 sections that overlap to form eight contigs. We numbered each section arbitrarily from 1 (nearest the telomere of the short arm) to 1,800 (nearest the long-arm telomere) for the purpose of this analysis (Table 3, 'ptel–qtel' numbers). We detected 220 high or medium stringency matches between a 9-Mb region of human 6p25.3 (segment 6.1, represented by accessions 3–97) and a region of about 9.9 Mb in a mouse BAC contig mapped to Mmu13A3 (Table 3 and Fig. 3). In addition, we found that the order of all 220 matches is consistent between the two maps, supporting the possibility that this is a conserved segment. The next segment (segment 6.2) starts in human accession 99 and extends as far as accession 207. This segment is defined by 303 matches of consistent order with a contig on Mmu13. Segments 6.1 and 6.2 are discontinuous in the mouse. The next segment (6.3) is aligned to another discontinuous region of Mmu13, whereas the fourth block (6.4) on human 6p is syntenic with a different mouse chromosome, Mmu17 (Table 3 and Fig. 3). In total, 20 conserved segments were identified along human chromosome 6, which subdivide the nine regions of conserved synteny (Fig. 3; see also Supplementary Information and updated views at http://www.ensembl.org/Homo_sapiens/synteniview). The results of this analysis can also be viewed with respect to the mouse chromosomes. For example, mouse chromosome 11 contains five regions of conserved synteny with the human genome, which are subdivided into 21 conserved segments (Fig. 4; see also Supplementary Information and updated views at http://www.ensembl.org/Mus_musculus/synteniview).

Overall, we defined 288 conserved segments contained within 167 regions of conserved synteny (Fig. 5). This compares with a previous

analysis using 2,997 markers, in which 183 conserved segments were detected⁶. The existence of multiple adjacent (but discontinuous) segments within the same syntenic region (for example segments 6.1–6.3 and 6.15–6.19 in Table 2 and Fig. 3) supports the hypothesis that a significant number of intrachromosomal inversions, in addition to interchromosomal rearrangements, contributed to formation of the present-day structures from the chromosomes of a common ancestor. Our results are consistent with previous observations for several chromosomal regions^{34–37}, including the recent high-resolution studies on human chromosomes 19 (refs 21, 38), 21 (ref. 39) and 7 (ref. 20). In the latter study, for example, comparative analysis using a high density of markers resulted in detection of 20 conserved segments within 16 regions of conserved synteny (compared with 25 and 15, respectively, in the present study). Small conserved segments will inevitably escape detection in the absence of sufficient homology crosslinks. For example, the Mmu2–Hsa7 segment reported in the previous study was estimated to be probably less than 100 kb (ref. 20), and was not detected in our study here. Anecdotal evidence from detailed comparisons of finished sequence between the mouse and human chromosomes indicates the presence of very small regions (as little as 1 kb) where conserved order is disrupted, either by recent duplications

or insertion of unrelated sequence^{23,34,39,40}. Ultimately, the alignment of finished sequences of the two genomes will be required to define the exact number and boundaries of every conserved segment.

Future work

Physical clone maps of complex genomes provide an essential framework to produce complete genome sequences. The clones also provide a source of well-characterized material for experimental studies. In the absence of large-scale sequencing programmes (and these may not be expected for a number of genomes in the foreseeable future), the availability of an accurate clone map will allow targeted sequencing of regions of specific interest. The alignment of maps of multiple genomes will also permit rapid investigation of orthologous gene sequences for structural and functional comparisons.

The first genomes to be mapped and sequenced in their entirety have been the result of enormous and expensive efforts. A hallmark of these studies has been the quality of the product, which has benefited from cross-validation between independently constructed maps for long-range information, and from the high accuracy of finished sequence. In our study, the availability of the human genome sequence has been exploited for faster characterization of the mouse genome. This approach will be applicable to many other mammals, and may also work for more distantly related vertebrates. Three factors contribute to the success of the mapping strategy: (1) good coverage of the genome in large contigs after the initial fingerprinting; (2) high-quality sequence cross-matches between most of these contigs and the reference sequence; and (3) sufficient independently mapped markers to order and orient all the contigs in the new map. Future projects would benefit from a similar depth of coverage in BACs (at least 15-fold) to ensure good representation of the genome in long contigs. Some savings might be made in the number of BESs required for the mapping alignment. One way to reduce the cost of the project would be to select clones from pre-existing contigs before sequencing the insert ends. However, this would reduce the number of sequence anchor points in the final map with consequent disadvantages (see below). For genomes more closely related than human and mouse, a higher proportion of the BESs would be expected to match the reference sequence (compared with 16% for the human–mouse study), which would allow reduction in the number of BESs generated without reduction of cross-matches.

The physical clone map of the mouse has provided 98% coverage of the genome in BAC contigs of average size more than 9 Mb. From the more than 300,000 BACs in the map, a complete tiling path of clones is being selected for production of the finished sequence. So

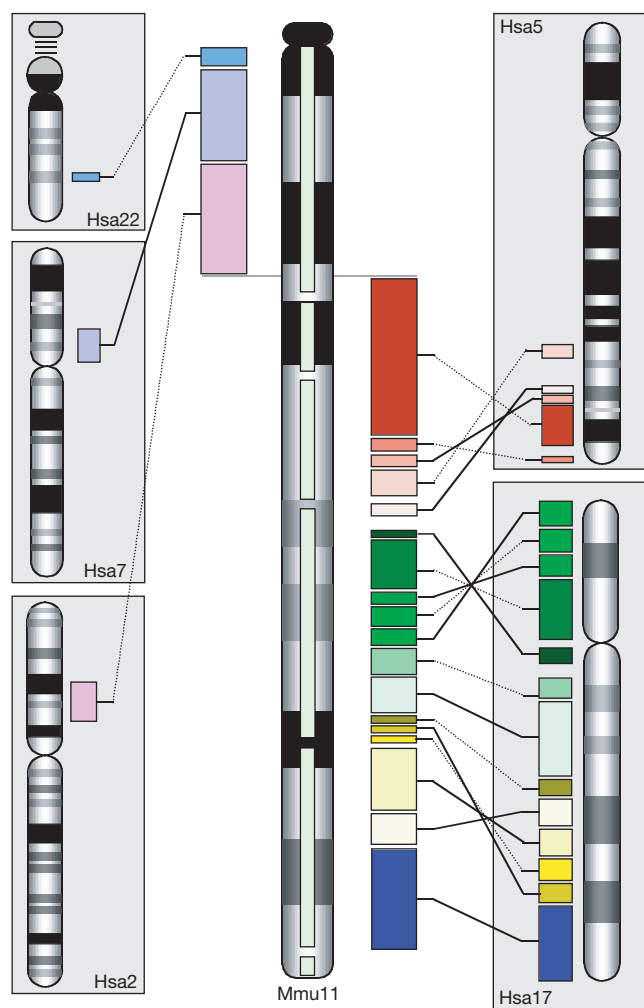


Figure 4 Conserved segments between mouse chromosome 11 and the human genome. The pale green bars in the centre of the chromosome ideogram represent the current extent of contig coverage in the mouse map. Dotted lines between the two species of same coloured blocks indicate conserved syntenic blocks inverted with respect to each other relative to the orientation of each chromosome as drawn.

Figure 5 Homology maps of the mouse and human genomes. **a**, The mouse–human homology map. Bars in the centre of each mouse chromosome ideogram denote the ordered and oriented mouse contigs determined by this study. The coloured blocks adjacent to each mouse chromosome indicate discrete regions where marker order is conserved between the two genomes. Colours correspond to specific human chromosomes in **b**. Positions of human telomeres are located within mouse contigs by arrows in the colour of their corresponding chromosome: up arrow indicates the p telomere and down arrow indicates the q telomere. The X-chromosome relationship is represented by arrows drawn in the direction Xp to Xq from the human. **b**, The human–mouse homology map. Bars in the centre of human chromosome ideograms denote clone contigs in the human genome. Adjacent coloured blocks indicate discrete regions where marker order is conserved with the mouse. Colours correspond to specific mouse chromosomes as in **a**. An interactive display of these figures is available (http://www.sanger.ac.uk/Projects/M_musculus/publications/fpcmap-2002). Regularly updated synteny information is also available (http://www.ensembl.org/Mus_musculus/syntenyview and http://www.ensembl.org/Homo_sapiens/syntenyview).

far, about 2.8% of the mouse genome sequence is available in this form (<http://www.ncbi.nlm.nih.gov>), and completion is scheduled for 2005. The mouse map contains more than 450,000 BESs (covering 220 Mb, or 7.8%, of the mouse genome), most of which contain sufficient unique sequence to anchor the current mouse whole genome sequence assemblies of 40 million reads ($\sim 6 \times$ coverage; <http://trace.ensembl.org>). This combined resource will provide a comprehensive working draft of the mouse genome in the public domain for immediate use. For example, the January 2002 release of this product included 20,652 annotated transcripts (http://www.ensembl.org/Mus_musculus/stats). The same sequence is also assisting the identification of genes on the basis of sequence conservation⁴¹ where no human cDNA is available. Detailed comparison between human and mouse sequences will therefore provide a comprehensive catalogue of genes (and orthologous-pair relationships where they exist) for comparative functional and genetic studies. It will also be possible to identify the evolutionary events at the sequence level that have occurred since divergence of the two species. □

Methods

Map construction

Fingerprints of RPCI-23 and RPCI-24 mouse BAC clones²² were generated by separation of *Hind*III restriction-enzyme digest fragments on 1.2% agarose gels⁴². Gel images were acquired with Image software (<http://www.sanger.ac.uk/Software/Image>) and restriction fragments identified using BANDLEADER software (D. Fuhrmann *et al.*, unpublished). Fingerprints were assembled within the program FPC⁴³ under high stringency conditions at a probability of 1×10^{-16} and a match tolerance of seven. The BESs⁴⁴ from clones within the fingerprint database were then matched to the tile path of human sequence clones. Human tile path accessions and mouse BESs were repeat masked and then aligned by BLASTN version 2.0a13MP-WashU, 10 June 1997 (<http://blast.wustl.edu>; W. Gish, personal communication)⁴⁵. All matches longer than 100 bp, and with a significance value above 0.01, were clustered using the initial fingerprint assembly. A 'high score' blast match (that is, with a blast score >700) was required to anchor each mouse map contig to the human tile path, supported by at least one other BLAST match (Fig. 1). Only matches with unique locations in the human tile path were used. Matches between the mouse BESs and human accessions had an average score of 627.29 and an average identity of 72.3%. Contigs that were adjacent to each other on the basis of their alignment to the human sequence were inspected manually and joined if fingerprints of the end clones overlapped with probability scores of better than 1×10^{-12} . Markers were added to the map either by electronic PCR³¹, or by hybridization using overgo probes³³. Markers were incorporated into the physical map only if they were uniquely placed within the map. Additional joins were made between contigs that were adjacent according to consistent mouse marker data if supported additionally by fingerprint overlaps with probability scores of better than 1×10^{-10} .

Coverage estimates

The exact length of 125 mouse clones, for which finished sequence was available (totalling 25,349 kb) was divided by the number of *Hind*III fragments (5,597) observed as single bands in the fingerprints of these clones. This provided a conversion factor of 4.5 kb per observed fragment, which was used to calculate all contig sizes in the consensus map as listed in Table 2. An estimate of 2.8 Gb for the size of the mouse genome was obtained as follows. Assembled mouse whole-genome shotgun data (November release, covering a total of 2.41 Gb) contained 91% coverage of 41 Mb of finished mouse genomic sequence. On this basis, the mouse genome is estimated to cover ~ 2.8 Gb (2.65 Gb euchromatin plus heterochromatin). Length estimates for individual mouse chromosomes were taken from refs 46 and 47, adjusted using the estimate of whole genome size described above (see Table 2). This result suggests that the mouse genome is shorter than the human genome. Support for this was obtained by comparing contiguous and finished mouse and human genomic sequence in 13 regions that could be aligned. The combined total lengths were 6.29 Mb (mouse) and 7.78 Mb (human) (ratio 0.8), although values for individual segments varied between 0.73 and 1.06. These results are in line with previous observations^{34,35,39,48–50}.

Received 15 April; accepted 4 July 2002; doi:10.1038/nature00957.

Published online 4 August 2002.

1. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).
2. Churcher, C. *et al.* The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX. *Nature* **387**, 84–87 (1997).
3. The yeast genome directory. *Nature* **387** (Suppl.), 5 (1997).
4. The *C. elegans* Sequencing Consortium Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
5. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
6. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

7. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
8. Collins, J. & Hohn, B. Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc. Natl Acad. Sci. USA* **75**, 4242–4246 (1978).
9. Shizuya, H. *et al.* Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA* **89**, 8794–8797 (1992).
10. Coulson, A., Sulston, J., Brenner, S. & Karn, J. Towards a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **83**, 7821–7825 (1986).
11. Olson, M. V. *et al.* Random-clone strategy for genomic restriction mapping in yeast. *Proc. Natl Acad. Sci. USA* **83**, 7826–7830 (1986).
12. Bentley, D. R. *et al.* The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X. *Nature* **409**, 942–943 (2001).
13. McPherson, J. D. *et al.* A physical map of the human genome. *Nature* **409**, 934–941 (2001).
14. Green, E. D. Strategies for the systematic sequencing of complex genomes. *Nature Rev. Genet.* **2**, 573–583 (2001).
15. Dunham, I. *et al.* The DNA sequence of human chromosome 22. *Nature* **402**, 489–495 (1999).
16. Waterston, R. & Sulston, J. E. The Human Genome Project: reaching the finish line. *Science* **282**, 53–54 (1998).
17. Nadeau, J. H. & Taylor, B. A. Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc. Natl Acad. Sci. USA* **81**, 814–818 (1984).
18. DeBry, R. W. & Seldin, M. F. Human/mouse homology relationships. *Genomics* **33**, 337–351 (1996).
19. Nadeau, J. H. & Sankoff, D. Counting on comparative maps. *Trends Genet.* **14**, 495–501 (1998).
20. Thomas, J. W. *et al.* Comparative genome mapping in the sequence-based era: early experience with human chromosome 7. *Genome Res.* **10**, 624–633 (2000).
21. Kim, J. *et al.* Homology-driven assembly of a sequence-ready mouse BAC contig map spanning regions related to the 46-Mb gene-rich euchromatic segments of human chromosome 19. *Genomics* **74**, 129–141 (2001).
22. Osogawa, K. *et al.* Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. *Genome Res.* **10**, 116–128 (2000).
23. Hardison, R. C., Oeltjen, J. & Miller, W. Long human–mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res.* **7**, 959–966 (1997).
24. Koop, B. F. *et al.* The human T-cell receptor TCRAC/TCRDC (C α /C δ) region: organization, sequence, and evolution of 97.6 kb of DNA. *Genomics* **19**, 478–493 (1994).
25. Ashworth, L. K. *et al.* An integrated metric physical map of human chromosome 19. *Nature Genet.* **11**, 422–427 (1995).
26. Kent, W. J. The BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
27. Copeland, N. G. *et al.* A genetic linkage map of the mouse: current applications and future prospects. *Science* **262**, 57–66 (1993).
28. Dietrich, W. F. *et al.* A comprehensive genetic map of the mouse genome. *Nature* **380**, 149–152 (1996).
29. Cai, W. W. *et al.* An SSLP marker-anchored BAC framework map of the mouse genome. *Nature Genet.* **29**, 133–134 (2001).
30. Hudson, T. J. *et al.* A radiation hybrid map of mouse genes. *Nature Genet.* **29**, 201–205 (2001).
31. Schuler, G. D. Sequence mapping by electronic PCR. *Genome Res.* **7**, 541–550 (1997).
32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
33. Ross, M. T., LaBrie, S., McPherson, J. P. & Stanton, V. P. In *Current Protocols in Human Genetics* (eds Dracopoli, N. C. *et al.*) 5.6.1–5.6.5 (Wiley, New York, 1999).
34. Puttagunta, R. *et al.* Comparative maps of human 19p13.3 and mouse chromosome 10 allow identification of sequences at evolutionary breakpoints. *Genome Res.* **10**, 1369–1380 (2000).
35. Carver, E. A. & Stubbs, L. Zooming in on the human–mouse comparative map: genome conservation re-examined on a high-resolution scale. *Genome Res.* **7**, 1123–1137 (1997).
36. Watkins-Chow, D. E. *et al.* Genetic mapping of 21 genes on mouse chromosome 11 reveals disruptions in linkage conservation with human chromosome 5. *Genomics* **40**, 114–122 (1997).
37. Stubbs, L. *et al.* Detailed comparative map of human chromosome 19q and related regions of the mouse genome. *Genomics* **35**, 499–508 (1996).
38. Dehal, P. *et al.* Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. *Science* **293**, 104–111 (2001).
39. Pletcher, M. T., Wiltshire, T., Cabin, D. E., Villanueva, M. & Reeves, R. H. Use of comparative physical and sequence mapping to annotate mouse chromosome 16 and human chromosome 21. *Genomics* **74**, 45–54 (2001).
40. Oeltjen, J. C. *et al.* Large-scale comparative sequence analysis of the human and murine Bruton's tyrosine kinase loci reveals conserved regulatory domains. *Genome Res.* **7**, 315–329 (1997).
41. Batzoglu, S., Pachter, L., Mesirov, J. P., Berger, B. & Lander, E. S. Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Res.* **10**, 950–958 (2000).
42. Marra, M. A. *et al.* High throughput fingerprint analysis of large-insert clones. *Genome Res.* **7**, 1072–1084 (1997).
43. Soderlund, C., Humphray, S., Dunham, A. & French, L. Contigs built with fingerprints, markers, and FPC V4. 7. *Genome Res.* **10**, 1772–1787 (2000).
44. Zhao, S. *et al.* Mouse BAC ends quality assessment and sequence analyses. *Genome Res.* **11**, 1736–1745 (2001).
45. Evans, S. F. & Gish, W. Local alignment statistics. *Methods Enzymol.* **266**, 460–480 (1996).
46. Nusbaum, C. *et al.* A YAC-based physical map of the mouse genome. *Nature Genet.* **22**, 388–393 (1999).
47. Evans, E. P. In *Genetic Variants of the Laboratory Mouse* (eds Lyon, M. F., Rastan, S. & Brown, S. D. M.) 1446 (Oxford Univ. Press, New York, 1996).
48. Pletcher, M. T. *et al.* Chromosome evolution: the junction of mammalian chromosomes in the formation of mouse chromosome 10. *Genome Res.* **10**, 1463–1467 (2000).
49. Martindale, D. W. *et al.* Comparative genomic sequence analysis of the Williams syndrome region (LIMK1-RFC2) of human chromosome 7q11.23. *Mamm. Genome* **11**, 890–898 (2000).
50. DeSilva, U. *et al.* Generation and comparative analysis of approximately 3.3 Mb of mouse genomic sequence orthologous to the region of human chromosome 7q11.23 implicated in Williams syndrome. *Genome Res.* **12**, 3–15 (2002).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>). Updated views of the map are available from the authors' websites (http://www.ensembl.org/Mus_musculus/cytoview and <http://www.ncbi.nlm.nih.gov/genome/guide/mouse>), as is an archive version for this publication, plus comparisons and discussion of genetic, RH and clone maps, and an interactive version of the synteny displays of Fig. 5 (http://www.sanger.ac.uk/Projects/M_musculus/publications/fpcmap-2002).

Acknowledgements

The authors acknowledge the support of the Wellcome Trust, the National Institutes of Health and the US Department of Energy. We are grateful to the web team at the Sanger

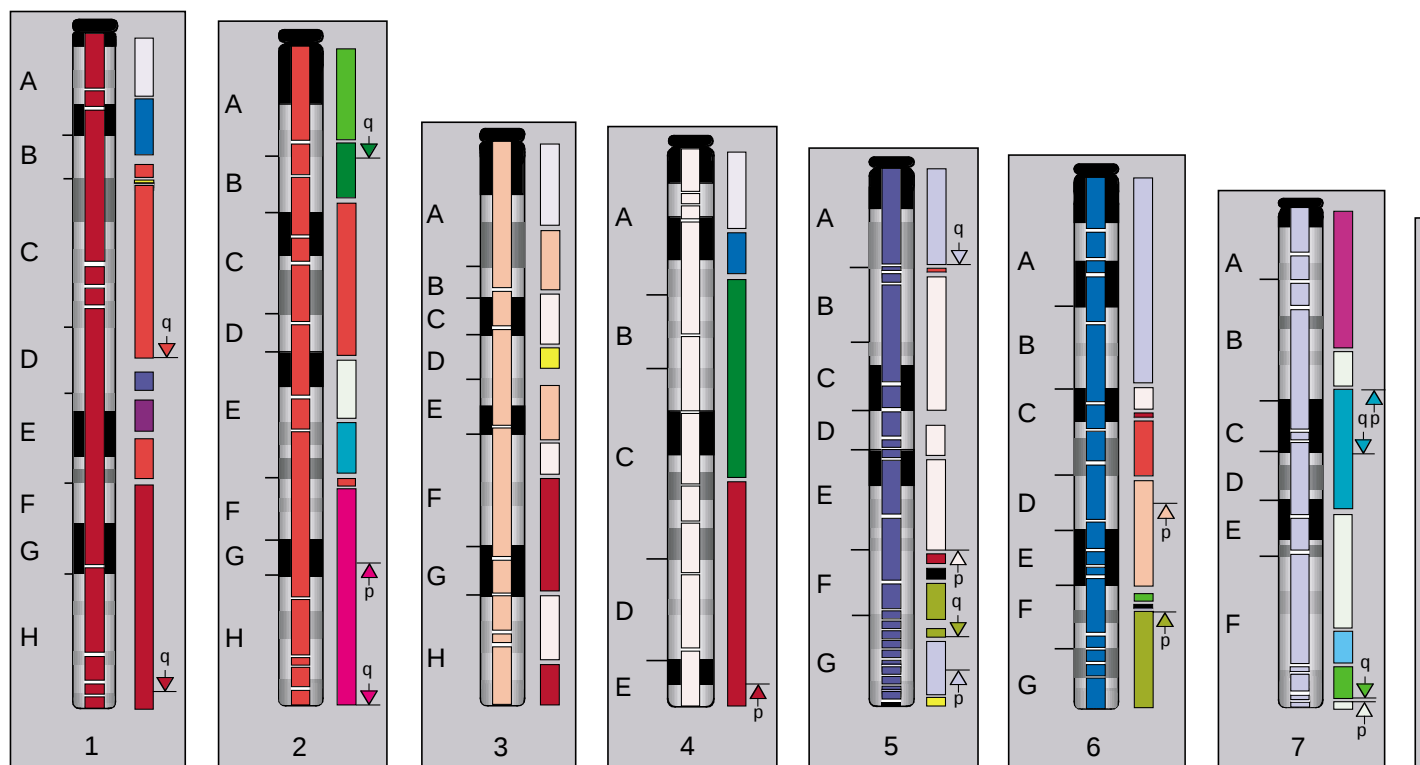
Institute for assistance with developing map displays, to P. Deloukas for RH map analysis, and E. Arnold-Berkowits, S. Lo, J. Gill and all present and past members of the Institute for Genomic Research BAC end sequencing team for the sequencing work.

Competing interests statement

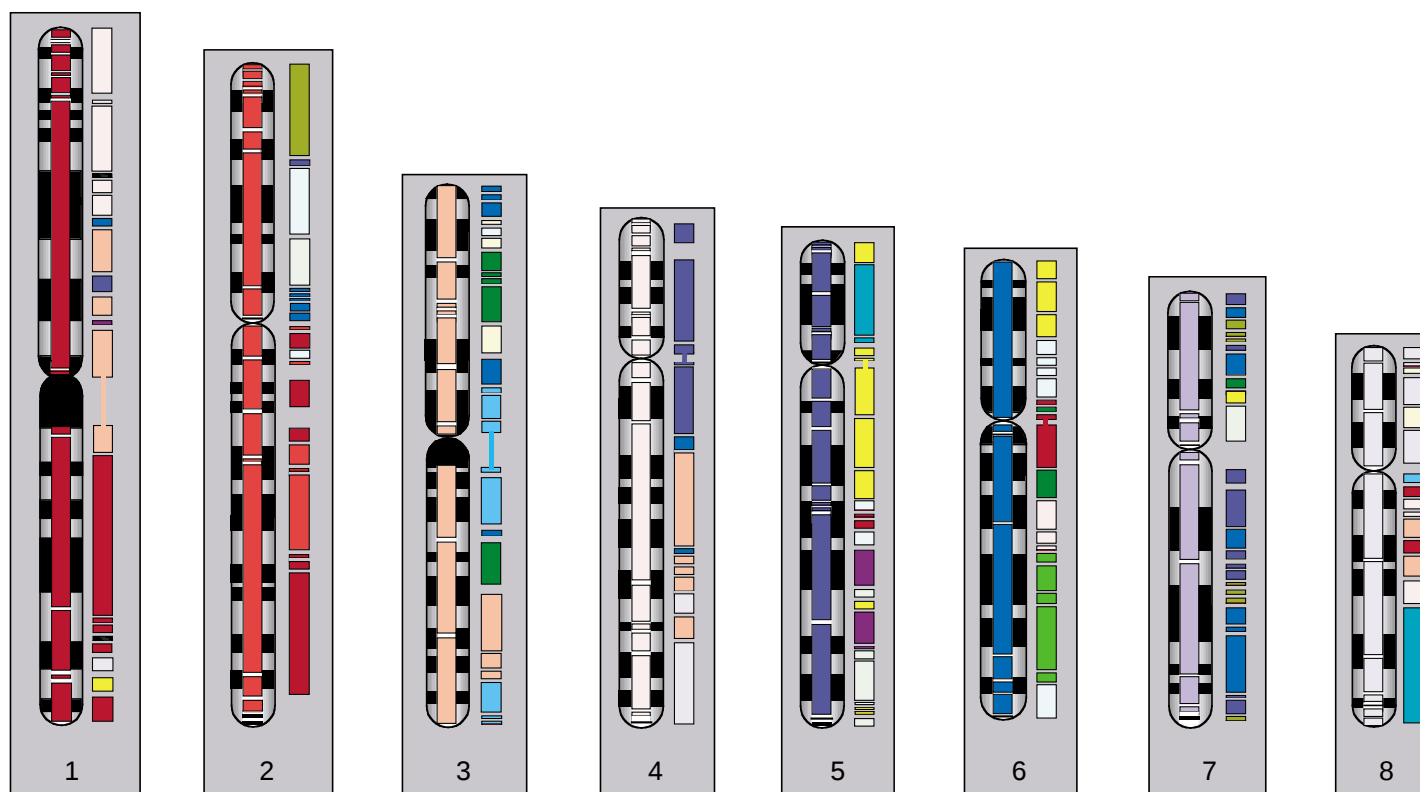
The authors declare that they have no competing financial interests.

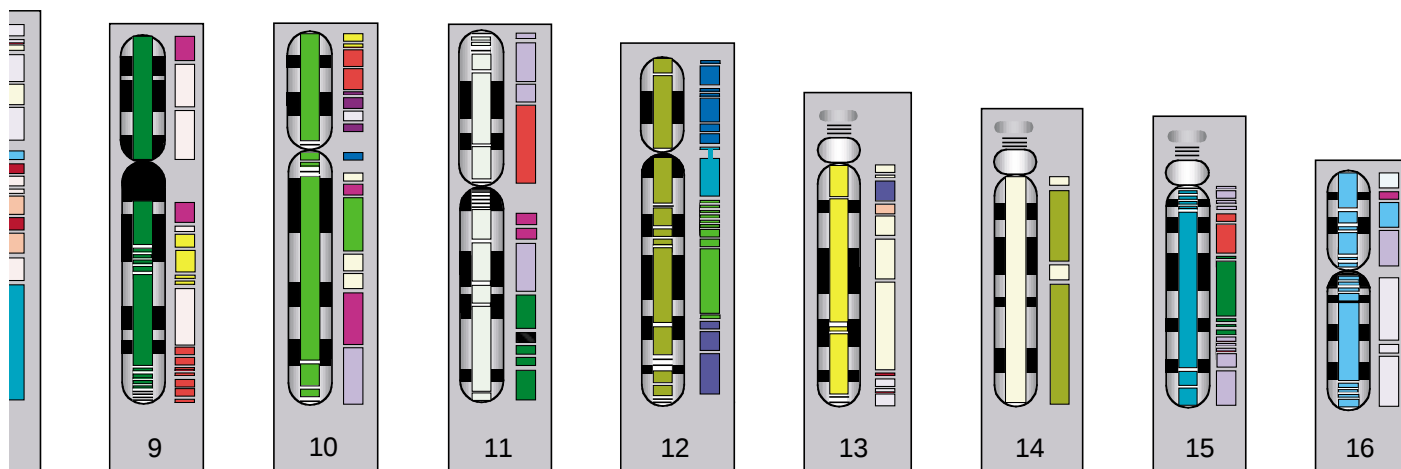
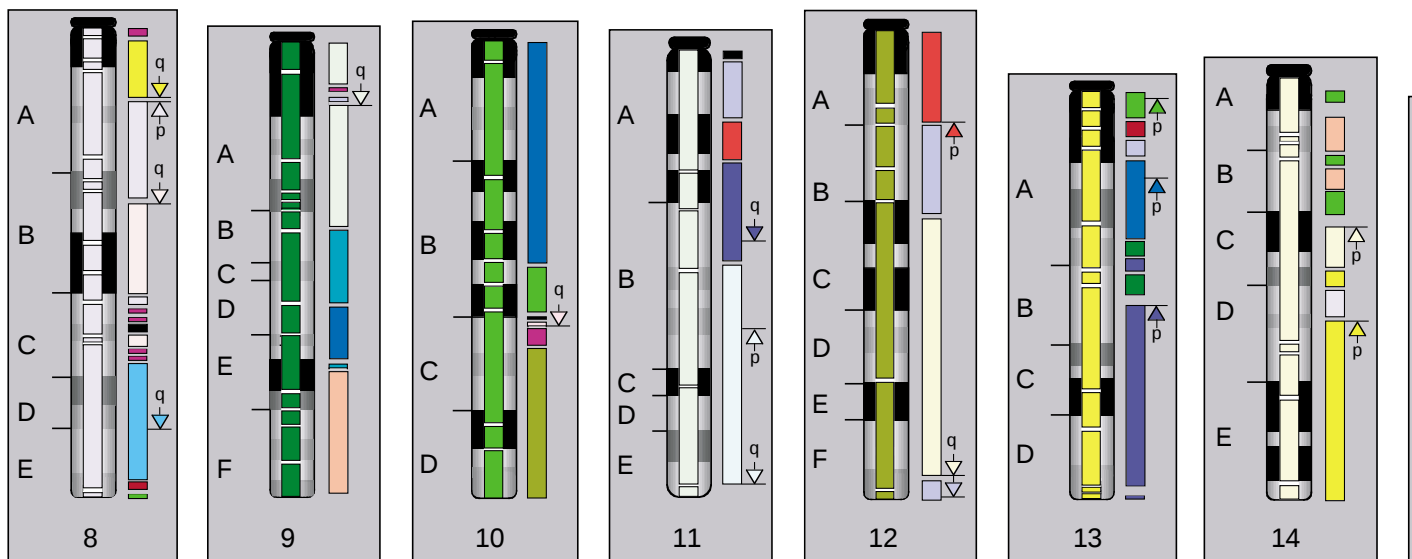
Correspondence and requests for materials should be addressed to D.R.B. (e-mail: drb@sanger.ac.uk).

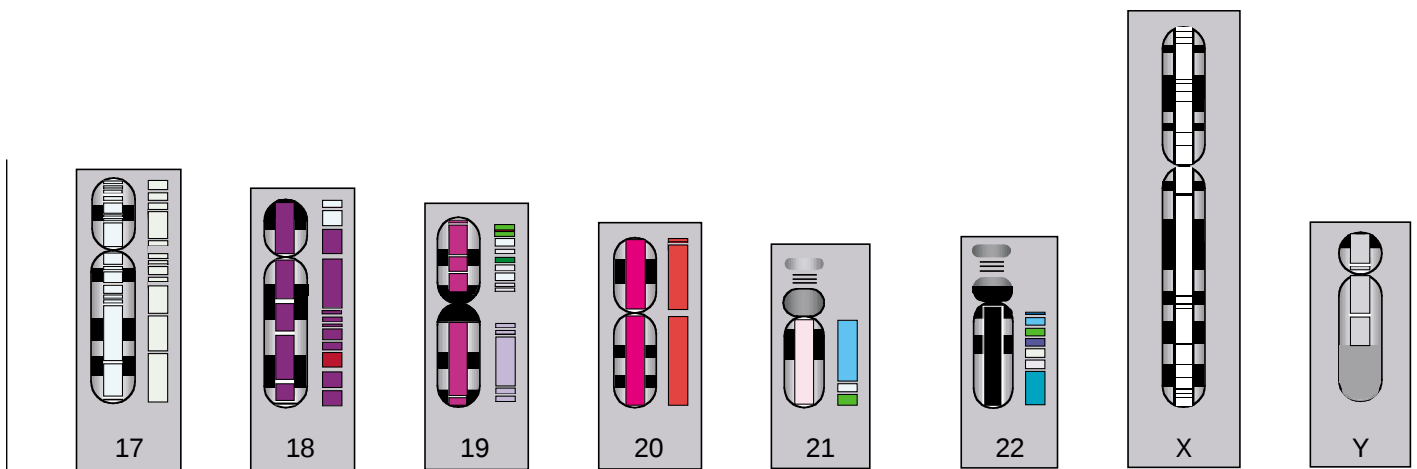
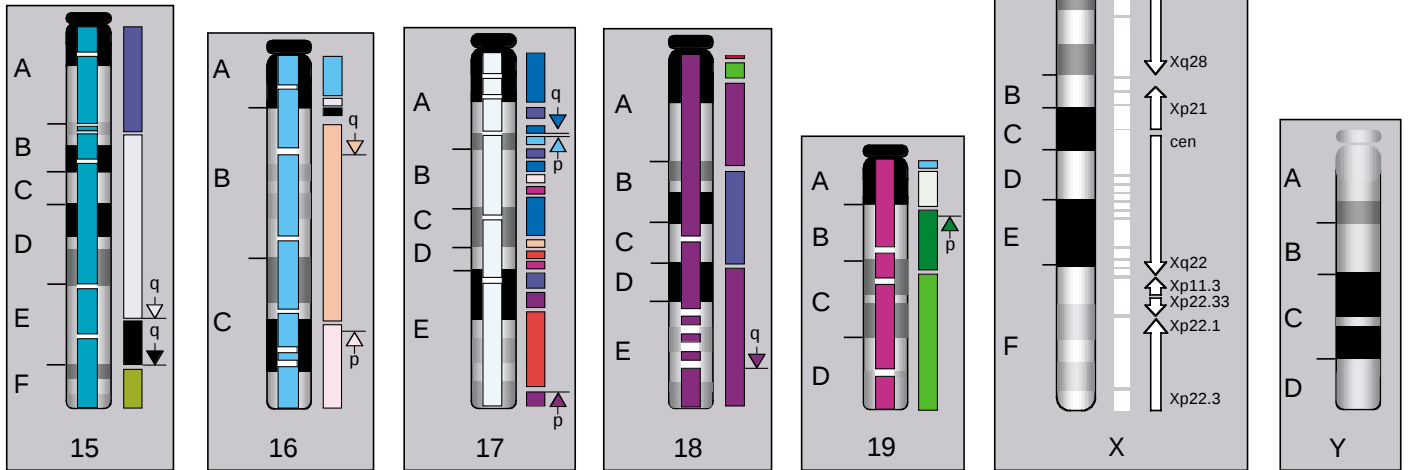
a Mouse-human homology map



b Human-mouse homology map







Macroscopically ordered state in an exciton system

L. V. Butov^{†‡}, A. C. Gossard[‡] & D. S. Chemla^{*§}

^{*} Materials Sciences Division, E. O. Lawrence Berkeley National Laboratory;
[§] Department of Physics, University of California at Berkeley, Berkeley,
 California 94720, USA

[†] Institute of Solid State Physics, Russian Academy of Sciences,
 142432 Chernogolovka, Russia

[‡] Department of Electrical and Computer Engineering, University of California,
 Santa Barbara, California 93106, USA

There is a rich variety of quantum liquids—such as superconductors, liquid helium and atom Bose–Einstein condensates—that exhibit macroscopic coherence in the form of ordered arrays of vortices^{1–4}. Experimental observation of a macroscopically ordered electronic state in semiconductors has, however, remained a challenging and relatively unexplored problem. A promising approach for the realization of such a state is to use excitons, bound pairs of electrons and holes that can form in semiconductor systems. At low densities, excitons are Bose-particles⁵, and at low temperatures, of the order of a few kelvin, excitons can form a quantum liquid—that is, a statistically degenerate Bose gas or even a Bose–Einstein condensate^{5–7}. Here we report photoluminescence measurements of a quasi-two-dimensional exciton gas in GaAs/AlGaAs coupled quantum wells and the observation of a macroscopically ordered exciton state. Our spatially resolved measurements reveal fragmentation of the ring-shaped emission pattern into circular structures that form periodic arrays over lengths up to 1 mm.

We studied spatially resolved photoluminescence (PL) of quasi-two-dimensional gases of indirect excitons in GaAs/AlGaAs coupled quantum wells (QWs); see Fig. 1b. Coupled QWs form a unique system where a cold exciton gas, and, more generally, a cold gas of light boson quasiparticles, can be created^{8–12}. The indirect excitons in coupled QWs are characterized by high cooling rates, three orders of magnitude higher than in bulk GaAs, and a long lifetime, more than three orders of magnitude longer than in a single GaAs QW. This lifetime is much longer than the characteristic timescale for cooling of initially hot photogenerated excitons down to temperatures well below 1 K, where the dilute Bose gas of indirect excitons becomes statistically degenerate¹⁰. Because the exciton mass, M , is small, smaller than the free electron mass m_0 , the quantum degeneracy temperature $T_0 = (\pi\hbar^2 n)/(2Mgk_B)$ (where g is the spin degeneracy of the exciton state, k_B is the Boltzmann constant and $\hbar$ is the Planck constant) exceeds 1 K at experimentally accessible exciton densities, n .

Another important advantage of the system is a repulsive interaction between the indirect excitons, characteristic of oriented dipoles: indirect excitons are dipoles oriented perpendicularly to the QW plane (Fig. 1b). The repulsive interaction stabilizes the exciton state against the formation of metallic electron–hole droplets^{13,14}, reinforces the Bose–Einstein condensation¹⁵ and results in a screening of an in-plane random potential (which is caused by interface fluctuations, impurities, and so on, and is unavoidable in any QW sample). In two-dimensional systems with a repulsive interaction a phase transition to a superfluid exciton state is possible at finite temperatures¹⁶. The latter is characterized by a long-range order at low temperatures¹⁷.

A high density of indirect excitons is achieved by nonresonant laser photoexcitation with energies at or above the direct exciton resonance where the photon absorption coefficient is high (Fig. 1b). In a quasiequilibrium, almost all photoexcited carriers relax to the indirect exciton states because they are lower in energy (the ratio

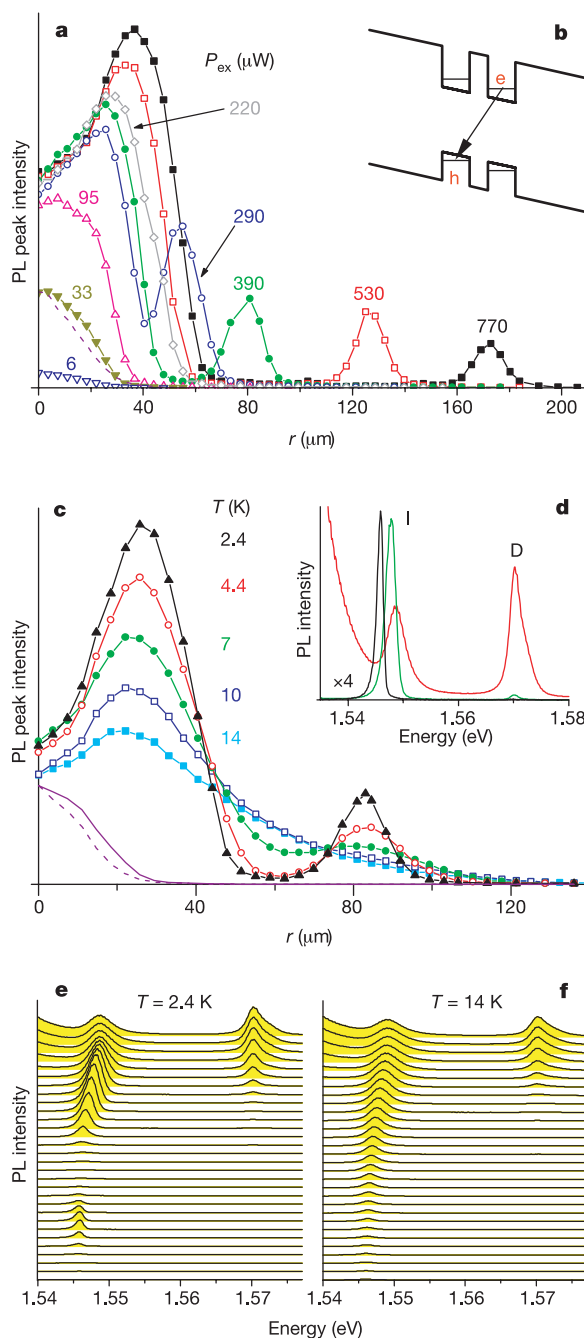


Figure 1 Radial dependence of the indirect exciton photoluminescence (PL). **a**, Peak intensity of the indirect exciton PL versus r , the distance from the excitation spot centre, at $T = 1.8$ K, gate voltage $V_g = 1.22$ V, and the excitation powers $P_{\text{ex}} = 6, 33, 95, 220, 290, 390, 530$ and $770 \mu\text{W}$. **c**, Peak intensity of the indirect exciton PL versus r at $P_{\text{ex}} = 390 \mu\text{W}$, $V_g = 1.22$ V, and $T = 2.4, 4.4, 7, 10$ and 14 K. The excitation spot profiles are shown by the purple dashed lines in **a** and **c**. The solid purple line in **c** shows the peak intensity of the direct exciton PL. The corresponding spatial dependence of the PL spectra at $T = 2.4$ and 14 K are shown in **e** and **f**. The uppermost spectra are recorded at $r = 0$, the lowest at $r = 107 \mu\text{m}$, and the step is $3.7 \mu\text{m}$. The indirect exciton PL line (I) is at about 1.545 – 1.55 eV; the direct PL line (D) is at about 1.57 eV; the broad line arising below the indirect exciton line is the bulk n^+ -GaAs emission. The selected spectra at $T = 2.4$ K are shown in **d**: in the excitation spot centre at $r = 0$ (red), in the internal ring centre at $r = 29 \mu\text{m}$ (green), and in the external ring centre at $r = 83 \mu\text{m}$ (black, the intensity is multiplied by 4). **b**, Energy band diagram of the CQW structures; e, electron; h, hole. The PL of indirect excitons is characterized by the rings centred at the excitation spot: the internal ring is located near the edge of the excitation spot, while the external ring is observed far away from the excitation spot.

between the indirect and direct exciton densities is typically $>10^4$). The initially photogenerated excitons are hot, but quickly cool down to the lattice temperature, T_{lattice} , by phonon emission. For example, the exciton temperature, T_X , can drop down to 400 mK in about 5 ns, that is, a time much shorter than the indirect exciton lifetime¹⁰. Therefore, there are two ways to overcome the obstacle of hot generation and study cold gases of indirect excitons with $T_X \approx T_{\text{lattice}}$: (1) use a separation in time and study the indirect excitons a few nanoseconds after the end of the photoexcitation pulses¹⁰; (2) use a separation in space and study the indirect excitons beyond the photoexcitation spot. In the second case, excitons can cool down to T_{lattice} as they travel away from the photoexcitation spot.

Here, exploring the spatially and spectrally resolved PL experiments, we have observed a ring structure in the indirect exciton PL and a macroscopically ordered state of indirect excitons appearing in the ring that is farthest from the excitation spot.

At the lowest excitation powers, P_{ex} , the spatial profile of the indirect exciton PL intensity closely follows the laser excitation intensity (Fig. 1a). However, at high P_{ex} we observed a nontrivial pattern for that profile. The pattern is characterized by a ring structure (Fig. 1): the laser excitation spot is surrounded by two

concentric bright rings separated by an annular dark intermediate region. The rest of the sample outside the external ring is dark. The internal ring appears near the edge of the laser excitation spot, and the external ring can be more than 100 μm from the excitation spot. Its radius increases with P_{ex} . The ring structure follows the laser excitation spot when it is moved over the whole sample area. This nontrivial spatial profile of the indirect exciton PL intensity is only observed at low temperatures. When the temperature is increased the bright rings wash out, the PL intensity in the inter-ring region and outside the external ring increases, and the profile approaches a monotonic bell-like shape (Fig. 1c).

The external ring is fragmented into circular structures that form a periodic array over macroscopic lengths, up to around 1 mm (Fig. 2a–e). This is demonstrated in Fig. 3e which shows the nearly linear dependence of the fragment positions along the ring versus their number. The fragments follow the external ring either when the excitation spot is moved over the sample area or when the ring

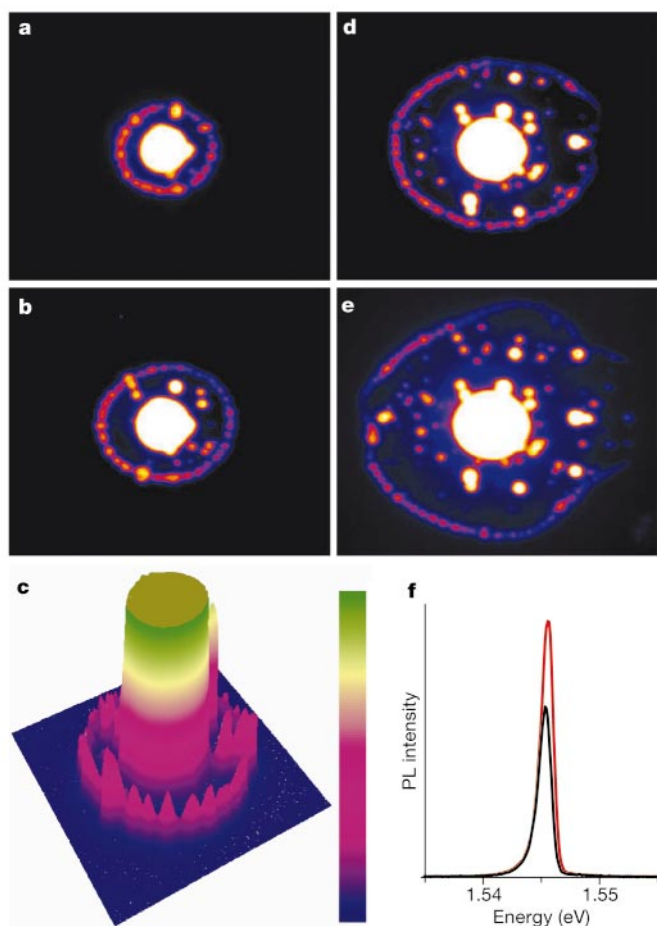


Figure 2 Excitation density dependence of the spatial pattern of the indirect exciton PL intensity. **a–e**, The pattern at $T = 1.8$ K, $V_g = 1.22$ V and $P_{\text{ex}} = 290$ (**a**), 390 (**b**), 690 (**d**), and 1,030 (**e**) μW . For **a**, **b**, **d** and **e** the area of view is $530 \times 440 \mu\text{m}$. The external ring of the indirect exciton PL is fragmented into a periodic chain of circular structures. The fragments follow the external ring both when its radius is changed by varying P_{ex} or when the laser excitation is moved over the sample area. The indirect exciton PL intensity is also strongly enhanced in some spots within the area terminated by the external ring. The position of these spots is fixed on the sample. **f**, Indirect exciton PL spectra in a peak (red) and the adjacent pass (black) on the fragment chain along the ring. Colour bar in **c** starts from zero (blue), and presents a linear scale in arbitrary units.

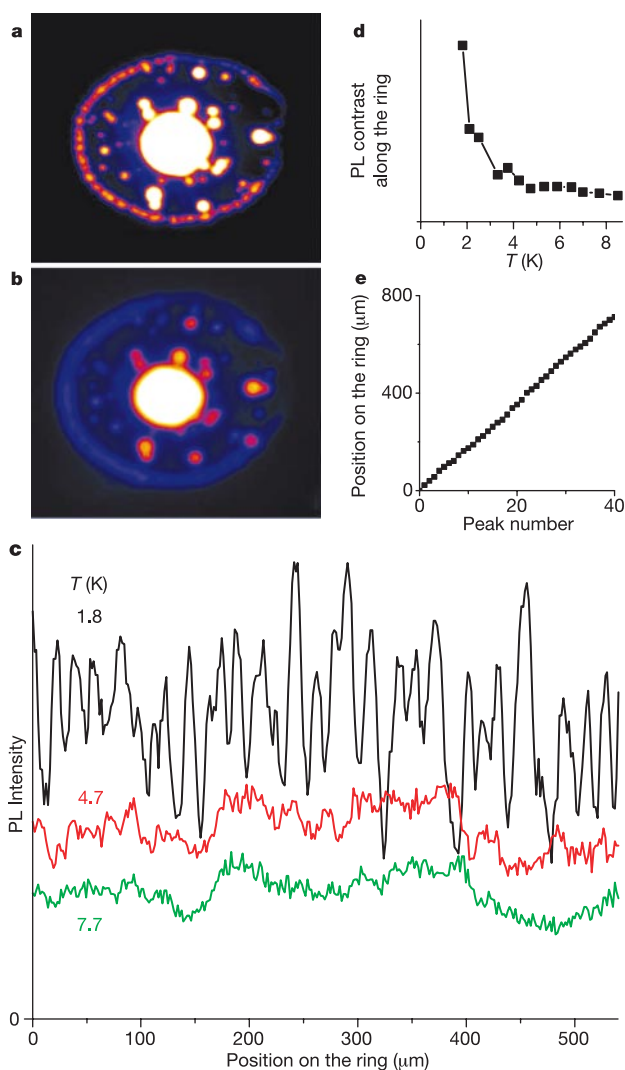


Figure 3 Temperature dependence of the spatial pattern of the indirect exciton PL intensity. **a**, **b**, The pattern at $T = 1.8$ (**a**) and 4.7 K (**b**) for $V_g = 1.22$ V, and $P_{\text{ex}} = 690 \mu\text{W}$. The area of view is $475 \times 414 \mu\text{m}$. **c**, The corresponding variation of the indirect exciton PL intensity along the external ring at $T = 1.8$, 4.7 and 7.7 K. The ring fragmentation into the periodic chain washes out with increasing temperature. This is visualized by the PL contrast (**d**) presented by an amplitude of the Fourier transform. The dependence of the position of the indirect exciton PL intensity peaks along the external ring versus the peak number is nearly linear (**e**), showing that the fragments form a periodic chain.

radius varies with P_{ex} . Along the whole external ring, both in the peaks and the passes, the indirect exciton PL lines are spectrally narrow with the full-width at half-maximum (FWHM) ≈ 1.3 meV, considerably smaller than in the centre of the excitation spot; see Figs 2f and 1d. The ring fragmentation is observed at the lowest temperatures only, $T \lesssim 4$ K (Fig. 3a–c). With increasing temperature, the PL contrast along the ring washes out; this is quantified by the amplitude of the Fourier transform of the PL intensity variation along the ring (Fig. 3d).

The spatial pattern shows also that the indirect exciton PL intensity is strongly enhanced in certain fixed spots on the sample; see Fig. 2a–e. We call them localized bright spots (LBS). For any excitation spot location and any P_{ex} the LBS are only observed when they are within the area terminated by the external ring (Fig. 2a–e). In the LBS the indirect exciton PL line is spectrally narrow, the FWHM ≈ 1.2 meV, and its energy is locally reduced. The LBS also wash out with increasing temperature.

All these effects are observed on several mesas studied and all experimental data are reproducible after cycling the sample temperature up to room temperature and back to 1.8 K many times. We now discuss the nature of the observed effects.

Under continuous-wave laser excitation a system of photoexcited excitons freely expanding in a QW plane is a thermodynamically open system in a quasiequilibrium. At low densities the indirect excitons are localized by the in-plane potential fluctuations and, therefore, the spatial profile of the indirect exciton PL intensity closely follows the profile of the laser excitation intensity (Fig. 1a). At high densities, however, the repulsive interaction between the indirect excitons screens the in-plane potential fluctuations and, therefore, results in a delocalization of the indirect excitons. Their long lifetime allows them to move far away from the excitation spot before they recombine. As we show below, the lifetime is further enhanced by the exciton motion itself. All these factors facilitate the exciton transport over macroscopic distances up to around 1 mm in our experiments.

For delocalized quasi-two-dimensional excitons only the states inside the radiative zone terminated by the photon cone (Fig. 4a) can recombine radiatively by resonant emission of photons¹⁸. Those are the states with small in-plane centre of mass momenta $K_{\parallel} \leq K_0 = E_g \sqrt{\epsilon}/(\hbar c)$ (E_g is the bandgap and ϵ is the dielectric constant). The exciton radiative decay rate and thus the exciton PL intensity

are determined by the fraction of excitons inside the radiative zone. In the centre of the excitation spot the exciton gas is characterized by a high T_X , larger than T_{lattice} (ref. 10). Under continuous-wave photoexcitation, there is a continuous flow of excitons out of the excitation spot owing to the exciton drift and diffusion (other mechanisms such as ballistic transport and phonon wind may also be contributing to the exciton cloud expansion). The exciton diffusion originates directly from the exciton density gradient. The exciton drift also originates from the density gradient as the latter gives rise to the gradient of the indirect exciton potential energy because of the repulsive interaction (see Fig. 1e, f). As the excitons travel away from the excitation spot, T_X decreases with increasing radial distance owing to the energy relaxation of the excitons. The reduction of T_X increases the radiative zone occupation and, therefore, increases the PL intensity; this is seen as the onset of the internal ring. The internal ring is therefore the spatial analogue of the PL jump observed in ref 10. Estimates for the exciton density in the internal ring exceed $3 \times 10^{10} \text{ cm}^{-2}$ at the highest P_{ex} , which at $T = 2$ K implies a statistically degenerate Bose gas of indirect excitons (at $n = 3 \times 10^{10} \text{ cm}^{-2}$ and $T = 2$ K the Bose–Einstein distribution function gives the occupation number of the lowest energy state $\nu = e^{T_0/T} - 1 \approx 0.3$ for the excitons in our coupled QWs where $g = 4$ and $M = 0.21m_0$).

To explain the dark region between the rings we propose the following scenario (Fig. 4). Travelling out of the potential energy ‘hill’ at the centre of the excitation spot excitons acquire an average drift momentum, K_{drift} . As the height of the potential energy hill at several millielectronvolts (Fig. 1e) is much larger than the kinetic energy at the radiative zone edge ($\hbar^2 K_0^2/(2M) \approx 0.1$ meV), K_{drift} can exceed K_0 . That means that the moving excitons become optically inactive. This explains the existence of the dark region between the internal and external rings. One aspect of this effect is that the optically inactive excitons move at speeds faster than the speed of sound, $v_s = 3.7 \times 10^5 \text{ cm s}^{-1}$: even at $K = K_0$, the speed of the excitons is $v = \hbar K_0/M = 1.4 \times 10^6 \text{ cm s}^{-1}$. Far from the excitation spot the main driving force for exciton transport, the energy gradient, vanishes and the excitons relax down to the lowest energy states. This results in the sharp enhancement of the radiative zone occupation and thus the PL intensity, that is seen as the external ring. We can only give a crude estimate for the exciton density in the external ring: $n \lesssim 10^{10} \text{ cm}^{-2}$. As excitons in the external ring relax down to the low momentum states, characterized by low velocities, the exciton flow is stopped and there are almost no excitons outside that ring (Fig. 2a–e).

The most interesting feature of the external ring is its fragmentation into a periodic array. The existence of this periodic ordering shows that the exciton state formed in the external ring has a coherence on a macroscopic length scale. The coherence is not driven by a laser excitation because in our experiment the photoexcited carriers experience many inelastic scatterings before the optically active indirect excitons are formed. Instead, the coherence spontaneously appears in the exciton system. We note that the macroscopic ordering is observed in the same temperature range as bosonic stimulation of exciton scattering¹⁰, below a few kelvin (compare Fig. 3d and Fig. 1f of ref. 10).

The microscopic nature of this ordered exciton state warrants future studies. Comparing with known phenomena, one may suggest that the fragments are vortices in the exciton system and that the ordering is the consequence of a repulsive interaction between the vortices. Within this model, the vortex rotation is continuously supported by the exciton flow out of the excitation spot. Similar to arrays of vortices in atom Bose–Einstein condensates⁴, the rotation can be initiated even without an apparent external torque (possibly by a small deviation from axial symmetry of the exciton flow due to the in-plane potential fluctuations, which could cause a branching of the exciton flow, perhaps similar to the branching observed for the electron flow¹⁹). We also note that a

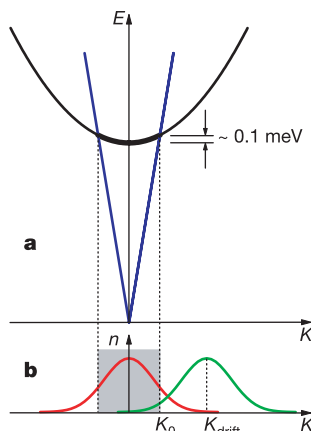


Figure 4 Schematics demonstrating a reduction of emission intensity for excitons in motion. **a**, Energy diagram for the exciton (black) and photon (blue) dispersion, the bold sector of the exciton dispersion indicates the radiative zone. **b**, The schematic momentum distribution of excitons without (red) and with (green) average drift velocity. The exciton radiative decay rate is proportional to the fraction of excitons in the radiative zone (grey area). E , energy; n , exciton density; K_0 , momentum at the radiative zone edge; K_{drift} , average drift momentum. See text.

spontaneous macroscopic flow organization with periodic vortical structures is a general property of thermodynamically open systems, both quantum^{1–4} and classical²⁰, described by nonlinear partial differential equations. A soliton array is another known phenomenon that may be related to the macroscopically ordered exciton state. Multiple solitons ('soliton trains') are known to be stationary solutions of the nonlinear Schrödinger equation on a ring (see ref. 21 and references therein). As the fragmentation that we observe appears abruptly at low temperatures, its nature is probably non-classical.

In-plane potential fluctuations could also influence the position of the fragments on the external ring owing to the pinning effect. However, the pinning effect appears to be small so that fragments remain free to move (for example, with changing of the excitation spot location on the sample) and the deviation of the fragment position out of the periodic array due to the pinning force is also small (Fig. 3e). In contrast, potential fluctuations can play a crucial role in formation of the LBS. We suggest that in the LBS the optically inactive moving indirect excitons are captured by a potential trap formed by in-plane potential fluctuations, relax to the optically active exciton states, and recombine, giving rise to the LBS emission. The LBS will be studied in detail later. In the context of the model discussed here, we note: (1) the existence of LBS in the dark annular region between the two rings confirms the presence of the optically inactive excitons in this region; and (2) the absence of LBS outside of the external ring confirms the absence of excitons there.

The spatial profiles of the direct exciton PL intensity exhibit none of the above effects and almost follow the laser excitation profile for all temperatures and P_{ex} studied; see Fig. 1c (this is also the case for the bulk GaAs emission). This is consistent with a short lifetime of direct excitons that limits the distance they can travel before the recombination and does not allow an effective cooling for them. □

Methods

The $n^+ - i - n^+$ GaAs/AlGaAs coupled QW structure was grown by MBE. The i -region consists of two 8-nm GaAs QWs separated by a 4-nm $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ barrier and surrounded by two 200-nm $\text{Al}_{0.33}\text{Ga}_{0.67}\text{As}$ barrier layers. The electric field in the sample growth direction is monitored by the external gate voltage V_g applied between the highly conducting n^+ -layers. For non-zero V_g the ground state is an indirect exciton made of an electron and a hole in different layers (Fig. 1b). The indirect exciton lifetime is in the range of tens and hundreds of nanoseconds. Radiative recombination is the dominant decay mechanism of indirect excitons in our high-quality sample. The indirect exciton energy shift with density, $\delta E(n)$, allows us to evaluate their concentration using $\delta E(n) = 4\pi n e^2 d / \epsilon$, where d is the effective separation between the electron and hole layers. The sample was excited with a HeNe laser at $\lambda = 632.8$ nm. Spatially resolved PL spectra were detected using a pinhole in the intermediate image plane. In the image experiment, the spatial pattern of the indirect exciton PL intensity was detected by charge-coupled device (CCD) camera using spectral selection of the indirect exciton emission by the interference filter. The spatial resolution was 5 μm .

Received 22 April; accepted 27 June 2002; doi:10.1038/nature00943.

- Essmann, U. & Trüble, H. The direct observation of individual flux lines in type II superconductors. *Phys. Lett. A* **24**, 526–527 (1967).
- Yarmchuk, E. J., Gordon, M. J. V. & Packard, R. E. Observation of stationary vortex arrays in rotating superfluid helium. *Phys. Rev. Lett.* **43**, 214–217 (1979).
- Madison, K. W., Chevy, F., Wohlleben, W. & Dalibard, J. Vortex formation in stirred Bose-Einstein condensate. *Phys. Rev. Lett.* **84**, 806–809 (2000).
- Abo-Shaeer, J. R., Raman, C., Vogels, J. M. & Ketterle, W. Observation of vortex lattices in Bose-Einstein condensates. *Science* **292**, 476–479 (2001).
- Keldysh, L. V. & Kozlov, A. N. Collective properties of excitons in semiconductors. *Sov. Phys. JETP* **27**, 521–528 (1968).
- Lozovik, Yu. E. & Yudson, V. I. A new mechanism for superconductivity: pairing between spatially separated electrons and holes. *Sov. Phys. JETP* **44**, 389–397 (1976).
- Perakis, I. E. Exciton developments. *Nature* **417**, 33–35 (2002).
- Fukuzawa, T., Mendez, E. E. & Hong, J. M. Phase transition of an exciton system in GaAs coupled quantum wells. *Phys. Rev. Lett.* **64**, 3066–3069 (1990).
- Butov, L. V. & Filin, A. I. Anomalous transport and luminescence of indirect excitons in AlAs/GaAs coupled quantum wells as evidence for exciton condensation. *Phys. Rev. B* **58**, 1980–2000 (1998).
- Butov, L. V. *et al.* Stimulated scattering of indirect excitons in coupled quantum wells: Signature of a degenerate Bose-gas of excitons. *Phys. Rev. Lett.* **86**, 5608–5611 (2001).
- Butov, L. V., Lai, C. W., Ivanov, A. L., Gossard, A. C. & Chemla, D. S. Towards Bose-Einstein condensation of excitons in potential traps. *Nature* **417**, 47–52 (2002).
- Larionov, A. V., Timofeev, V. B., Hvam, J. & Soerensen, K. Collective state of interwell excitons in GaAs/AlGaAs double quantum wells under pulse resonance excitation. *JETP Lett.* **75**, 200–204 (2002).

- Yoshioka, D. & MacDonald, A. H. Double quantum well electron-hole systems in strong magnetic fields. *J. Phys. Soc. Jpn* **59**, 4211–4214 (1990).
- Zhu, X., Littlewood, P. B., Hybertsen, M. & Rice, T. Exciton condensate in semiconductor quantum well structures. *Phys. Rev. Lett.* **74**, 1633–1636 (1995).
- Leggett, A. J. Bose-Einstein condensation in the alkali gases: Some fundamental concepts. *Rev. Mod. Phys.* **73**, 307–356 (2001).
- Popov, V. N. On the theory of the superfluidity of two- and one-dimensional Bose systems. *Theor. Math. Phys.* **11**, 565–573 (1972).
- Kosterlitz, J. M. & Thouless, D. J. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C* **6**, 1181–1203 (1973).
- Feldmann, J. *et al.* Linewidth dependence of radiative exciton lifetimes in quantum wells. *Phys. Rev. Lett.* **59**, 2337–2340 (1987).
- Topinka, M. A. *et al.* Coherent branched flow in a two-dimensional electron gas. *Nature* **410**, 183–186 (2001).
- Taylor, G. I. Stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. R. Soc. Lond. A* **223**, 289–343 (1923).
- Carr, L. D., Clark, C. W. & Reinhardt W. P. Stationary solutions of the one-dimensional nonlinear Schrödinger equation. I. Case of repulsive nonlinearity. *Phys. Rev. A* **62**, 063610–1–063610-10 (2000).

Acknowledgements

We thank A.L. Ivanov for discussions, C.W. Lai and A.V. Mintsev for help in preparing the experiment and K.L. Campman for growing the high quality CQW samples. This work was supported by the Office of Basic Energy Sciences US Department of Energy and by the Russian Foundation for Basic Research (RFBR).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.V.B. (e-mail: lvbutov@lbl.gov).

Long-range transport in excitonic dark states in coupled quantum wells

D. Snoke*, S. Denev†, Y. Liu*, L. Pfeiffer† & K. West†

* Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pennsylvania 15260, USA

† Bell Labs, Lucent Technologies, 700 Mountain Avenue, Murray Hill, New Jersey 07974-0636, USA

During the past ten years, coupled quantum wells have emerged as a promising system for experiments on Bose condensation of excitons, with numerous theoretical^{1–6} and experimental^{7–12} studies aimed at the demonstration of this effect. One of the issues driving these studies is the possibility of long-range coherent transport of excitons. Excitons in quantum wells typically diffuse only a few micrometres from the spot where they are generated by a laser pulse; their diffusion is limited by their lifetime (typically a few nanoseconds) and by scattering due to disorder in the well structure. Here we report photoluminescence measurements of InGaAs quantum wells and the observation of an effect by which luminescence from excitons appears hundreds of micrometres away from the laser excitation spot. This luminescence appears as a ring around the laser spot; almost none appears in the region between the laser spot and the ring. This implies that the excitons must travel in a dark state until they reach some critical distance, at which they collectively revert to luminescing states. It is unclear whether this effect is related to macroscopic coherence caused by Bose condensation of excitons.

Coupled quantum wells are attractive for Bose condensation of excitons because the lifetime of the excitons can be extended by application of an electric field. A typical band structure is shown in Fig. 1. When electric field is applied normal to the planes of the wells, electrons are confined to one well and holes are confined in

the adjacent well. Because the electron and hole wavefunctions in the two wells have little overlap, excitons in this type of structure have much longer lifetimes, up to 100 ns (ref. 13). Excitons formed from an electron in one well and a hole in the adjacent well are known as 'indirect' excitons. The electric field shifts the electron and hole states to lower energy, in what is known as the 'quantum confined Stark effect'^{14,15}, so it is easy to identify the indirect exciton luminescence, because only the indirect exciton luminescence shifts to lower energy with increasing electric field.

Our present studies have focused on InGaAs coupled wells which have a confined state below the GaAs substrate on which they are grown. When this system is illuminated with red light (630 nm) well above the bandgap of the GaAs substrate, carriers generated in the substrate will fall into excitonic states in the coupled quantum wells, which are the lowest excited state in the system. This is therefore a very efficient way of creating a high density of excitons in the quantum wells. At very high excitation density, the carriers screen out the electric field, leading to a blue shift of the exciton luminescence as seen in ref. 16, but there exists a regime of density in which the indirect exciton luminescence does not shift substantially with excitation density.

In this regime we find that above a critical excitation density, the indirect exciton luminescence exhibits a ring structure centred on the laser spot. Figure 2 shows typical images from our experimental data. An image of the sample is projected onto the entrance slot of an imaging spectrometer, which preserves the spatial information along the slit while spectrally dispersing the light in the perpendicular direction. These images from an intensified charge-coupled device (CCD) camera on the back of the spectrometer therefore give the spectrum of the exciton luminescence as a function of position. The ring appears as two spots on these images. As seen in Fig. 2, the radius of the ring depends on the laser excitation density. The radius of the ring can also be changed by keeping the laser intensity constant and adjusting the applied voltage across the wells.

As seen in these images, the ring appears many hundreds of micrometres distant from the laser spot, but the intervening region is dark. We have confirmed that the structure is a ring centred on the laser spot by moving the laser spot; when the laser spot is moved upward, the ring follows the laser, remaining centred on it, and when the laser spot is moved to the side, only the two spots from the ring structure are seen in the CCD images, with decreasing separation until they merge when the edge of the ring is projected onto the entrance slit of the spectrometer.

The most surprising aspect of these images is that no luminescence appears in the space between the laser spot and the ring for hundreds of micrometres. This implies that the excitons must travel in some dark state until they reach some critical distance, at which they collectively revert to luminescing states. This type of behaviour is unprecedented in other excitonic systems.

On one hand, it would seem natural to interpret this effect as arising from Bose coherent effects of the excitons, but since this

effect would be a retrodiction, not a prediction, of any present theory of Bose coherence of excitons, we do well to ask what other tests show. One obvious test is to determine the temperature of the critical density. The critical temperature for Bose superfluidity depends on the density; alternatively, we can speak of the critical density at a given temperature. We have observed the ring structure at a series of different temperatures and determined the critical excitation density at which the ring appears for each temperature.

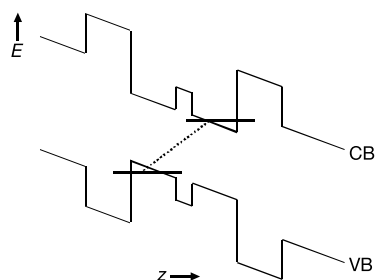


Figure 1 The band structure of the coupled quantum wells in the presence of electric field. Indirect excitons are made from an electron in one well and a hole in the adjacent well. (The dashed line indicates the optical transition which gives rise to the luminescence from indirect excitons.) *E*, energy; *z*, distance; CB, conduction band; VB, valence band.

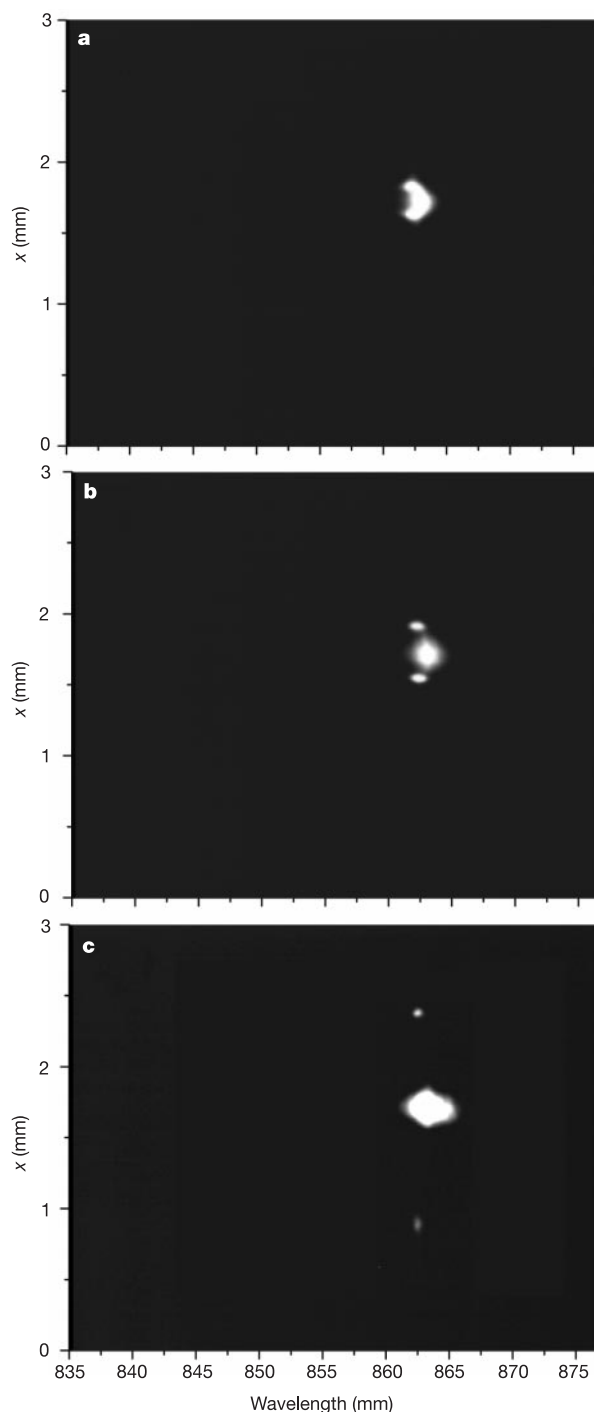


Figure 2 Spectral images of the luminescence from indirect excitons for various laser powers. **a**, 1.7 mW; **b**, 2.0 mW; **c**, 2.5 mW. The ring around the laser spot is seen in these figures as a pair of dots equidistant from the intense, central spot, which appears at the focal spot of the laser beam which creates the excitons. The horizontal axis gives the wavelength of the luminescence. The ring appears at slightly shorter wavelength (higher energy) than the luminescence from the excitons at the laser focus.

The excitation density is simply the number of photons per laser pulse divided by the area of the laser spot. Because the laser at this wavelength is essentially 100% absorbed by the substrate near the quantum wells, the excitation density will be approximately the same as the exciton density in the wells if the exciton lifetime is less than the period between laser pulses; if the exciton lifetime is much longer than this period, the exciton density can greatly exceed this value.

The results for the critical excitation density as a function of temperature are shown in Fig. 3. The critical density is defined by the point at which the ring radius is zero, as shown in the inset of Fig. 3. In our coupled quantum well structure, the binding energy of the indirect excitons is approximately 10 meV. At $T = 90$ K, free electrons and holes coexist with the excitons, but we still see the ring structure at high laser excitation density. No free electrons and holes are seen in the dark region between the central spot and the ring, which indicates that the transport occurs entirely in dark states even at these temperatures.

For a Kosterlitz–Thouless transition to a superfluid state in two dimensions, the critical density is expected to be directly proportional to the temperature¹⁷, equal to

$$\sigma_c = \frac{2k_B T m}{\pi \hbar^2} \quad (1)$$

which gives a critical density of $8 \times 10^{10} \text{ cm}^{-2}$ for an effective exciton mass of $m = 0.12m_0$ (the expected in-plane effective mass of excitons in InGaAs quantum wells) and $T = 10$ K. As seen in Fig. 3, the critical density is approximately proportional to the square of the temperature, instead of linear. This would tend to argue against the possibility of a Bose coherent transition. On the other hand, for Bose condensation in a potential minimum, the critical number for condensation is proportional to the square of the temperature¹⁷, and the critical number of excitons can be far less, depending on the properties of the potential minimum. Butov *et al.*¹⁰ have argued that indirect excitons in this type of coupled quantum well can be trapped in a local minimum created by random fluctuations in the quantum well composition. Alternatively, one could argue that many-body effects of the exciton gas create a potential minimum wherever the excitons are, for example a local reduction of the bandgap due to heating of the lattice. There is evidence for this in the fact that the indirect exciton luminescence at the laser spot always appears slightly lower in energy than the luminescence from the ring. As seen in Fig. 2, the indirect exciton luminescence at the laser spot is around 1 meV lower in energy. This

fact raises the question, however, of why the Bose-condensed excitons would leave the trap to form the ring at higher energy. The standard prediction for Bose condensation in a trap is that the condensate will form in the centre, at the point of lowest energy.

One possible mechanism for the outward expansion is that a ‘phonon wind’ pushes the excitons out of the central region. It has long been known^{18–20} that a laser with photon energy well above the bandgap can generate a nonequilibrium distribution of phonons that act like a wind to push excitons. Normally, however, such a wind does not create a well-defined ring structure in a time-integrated image; the excitons emit luminescence continuously as they move. A phonon wind should also be much less efficient at high temperature.

A second test is to ask whether time resolving the luminescence gives results consistent with migration of the excitons from the laser spot to the ring. At typical exciton velocities at low temperature on the order of micrometres per nanosecond, excitons should take hundreds of nanoseconds to travel hundreds of micrometres. We have performed time-resolved studies of the rings, using time-resolved single-photon counting with 40-ps resolution and a cavity-dumped, mode-locked pulsed laser with 2-ps pulse duration. As seen in Fig. 4a, the indirect exciton luminescence at the laser spot has a fairly short lifetime, around 3 ns, while the luminescence from the ring has a long lifetime. By changing the repetition rate of the laser using a cavity dumper to give very long periods between the pulses, we can directly measure the long-term behaviour of the excitons at the ring. Figure 4b shows the intensity of the ring as a function of time after a laser pulse, for three different distances of the ring from the central spot. The rise time of the luminescence is longer when the ring is further from the central spot, consistent with the greater distance travelled in dark states, and the exciton lifetime is also consistent with travel over such long distances. The rise times shown in Fig. 4b correspond to a velocity of the excitons of approximately $1.5 \times 10^6 \text{ cm s}^{-1}$, which is reasonable for drift-like motion of excitons but three times faster than the speed of sound in GaAs, which would seem to rule out a photon wind mechanism. These data also rule out the possibility of a radiative reabsorption effect, by which one could imagine that the ring occurs not by actual migration of the excitons, but by recombination into photons which are then reabsorbed further away in the well. Radiative reabsorption would give a rise time and lifetime the same as that of the luminescence in the central spot.

In general, there are two possibilities for dark states in GaAs quantum wells. Hot excitons will not emit luminescence if they have

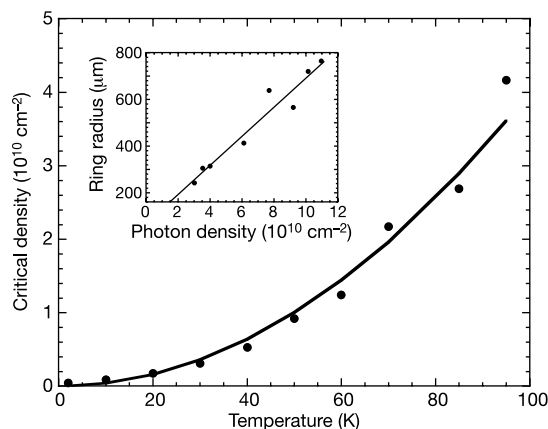


Figure 3 The critical excitation density for appearance of the luminescence ring, as a function of the bath temperature in a variable-temperature cryostat. The solid line is a fit to T^2 dependence. Inset, the ring radius versus excitation density. The critical threshold is

defined as the point at which the ring merges with the the central spot, which occurs in this case at $x = 160 \mu\text{m}$. The solid line is a linear fit of the data. At higher excitation density, the radius saturates and even decreases with increasing power.

energy greater than the linewidth of the direct recombination luminescence line, which is less than 1 meV. Excitons cool to the lattice temperature within a few nanoseconds, however, so it is unlikely that they all remain in high-energy states, with no emission from the low-energy states, for hundreds of nanoseconds. The excitons also have dark states which are nearly degenerate with the bright states²¹ but may lie slightly lower in energy. A condensate in one of these states could prevent luminescence until the excitons revert to the normal state.

It is worth asking what is different in these experiments from previous experiments, and why this ring has not been seen before in similar structures. There are two main differences in these new experiments. First, the quality of these new quantum well structures is very high. The full width at half maximum of the luminescence line at low temperature and low excitation density in these InGaAs quantum well samples ranges from 0.7 to 1.2 meV, which is very low compared to typical narrow quantum well structures. Second, many previous experiments have used near-resonant excitation, under the assumption that the exciton temperature would be lower with less excess photon energy. In these experiments, we use red (~630 nm) light which is well above the bandgap of the substrate. The absorption

of the laser light in this case is nearly 100%, compared to around 5% for near-resonant excitation. Most of the light is absorbed in the substrate, after which free carriers can fall into the quantum wells. It may be that producing a large number of free carriers in this way is an ideal method for achieving a high density of excitons; alternatively, the high-energy carriers may lead to unusual screening effects that have nothing to do with Bose coherence. As discussed above, the red laser light may also have the effect of creating a phonon wind and self trapping due to heating of the lattice. Finally, it is possible that this effect has occurred but not been observed before, because the luminescence from the ring structure is weak and would not be observed without an intensified imaging system.

Because of the unsettled issues discussed above, it is not clear whether this effect is related to Bose coherence. However, the transport of excitonic signal over such large distances, with a dark intermediate state, is an effect which deserves further study. With exciton motion over such long distances, it is possible to imagine 'excitonic circuits', in which signals are carried by excitonic degrees of freedom, similar to 'spintronics', in which signal is carried by spin. □

Received 16 May; accepted 1 July 2002; doi:10.1038/nature00940.

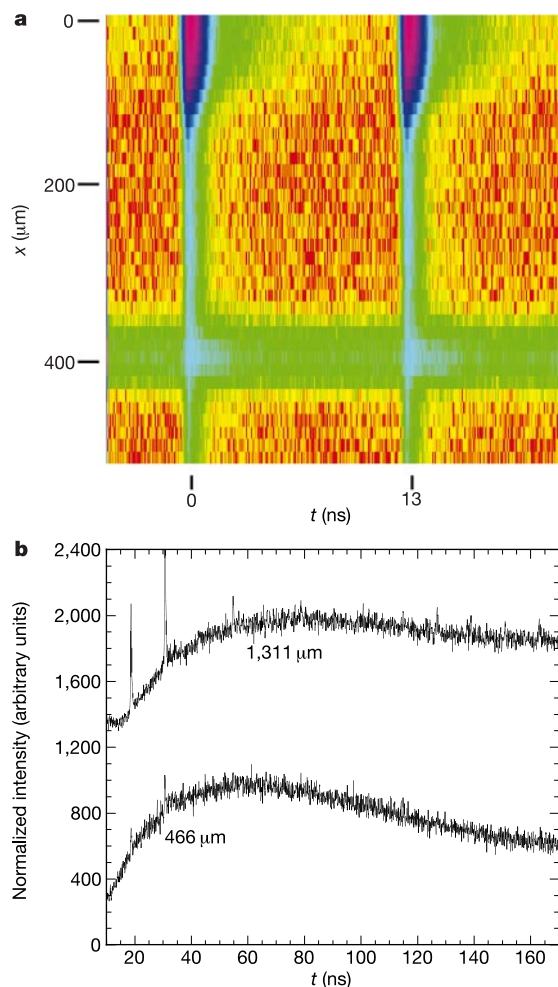


Figure 4 Luminescence profile and intensity. **a**, Time-resolved streak image of the indirect exciton luminescence profile. The centre of the laser spot, corresponding to $x = 0$, has short lifetime, while the ring (at $x = 400 \mu\text{m}$) has long lifetime. The laser repetition period in this case is 13 ns. No light appears in the region between the central spot and the ring at any time, except for scattered laser light during the laser pulses. **b**, Intensity of the luminescence from the ring as a function of time delay after the laser pulse, for two different ring radii, for a laser repetition period of 260 ns. The luminescence in the case of the $1,311\text{-}\mu\text{m}$ ring is offset for clarity.

- Zhu, X., Littlewood, P. B., Hybertson, M. S. & Rice, T. M. Exciton condensate in semiconductor quantum well structures. *Phys. Rev. Lett.* **74**, 1633–1636 (1995).
- Fernández-Rossier, J. & Tejedor, C. Spin degree of freedom in two-dimensional condensates. *Phys. Rev. Lett.* **78**, 4809–4812 (1997).
- Lozovik, Yu. E. & Birman, O. L. Phase transitions in a system of spatially separated electrons and holes. *JETP Lett.* **84**, 1027–1035 (1997).
- Yudson, V. I. Charged 'few-electron–single spatially separated hole' complexes in a double quantum well near a metal plate. *Phys. Rev. B* **77**, 1564–1567 (1996).
- Iida, T. & Tsubota, M. Order formation and superfluidity of excitons in type-II semiconductor quantum wells. *Phys. Rev. B* **60**, 5802–5810 (1999).
- Dzubenko, A. B. & Yablonskii, A. L. Intrawell and interwell magnetoexcitons in $\text{In}_x\text{Ga}_{1-x}$ coupled double quantum wells. *Phys. Rev. B* **53**, 16355–16364 (1996).
- Larionov, A. V. & Timofeev, V. B. Condensation of interwell excitons in GaAs/AlGaAs double quantum wells. *JETP Lett.* **73**, 342–350 (2001).
- Krivolapchuk, V. V., Moskalenko, E. S., Zhmodikov, A. L., Cheng, T. S. & Foxon, C. T. Collective properties of spatially indirect excitons in asymmetric GaAs/AlGaAs double quantum wells. *Solid State Commun.* **111**, 49–54 (1999).
- Butov, L. V. & Finin, A. I. Anomalous transport and luminescence of indirect excitons in AlAs/GaAs coupled quantum wells as evidence for exciton condensation. *Phys. Rev. B* **58**, 1980–2000 (1998).
- Butov, L. V., Lai, C. W., Ivanov, A. L., Gossard, A. C. & Chemla, D. S. Towards Bose–Einstein condensation of excitons in potential traps. *Nature* **417**, 47–52 (2002).
- Negoita, V., Snoke, D. W. & Eberl, K. Stretching quantum wells: a method for trapping free carriers in GaAs heterostructures. *Appl. Phys. Lett.* **75**, 2059–2061 (1999).
- Negoita, V., Hackworth, D., Snoke, D. W. & Eberl, K. Sub-Hz spectral fluctuations from high-density excitons in biased coupled quantum wells. *Opt. Lett.* **25**, 572–574 (2000).
- Negoita, V., Snoke, D. W. & Eberl, K. Harmonic potential traps for indirect excitons in coupled quantum wells. *Phys. Rev. B* **60**, 2661–2669 (1999).
- Fox, A. M., Miller, D. A. B., Livescu, G., Cunningham, J. E. & Yan, W. Y. Excitonic effects in coupled quantum wells. *Phys. Rev. B* **44**, 6231–6242 (1991).
- Kato, Y., Takahashi, Y., Fukatsu, S., Shiraki, Y. & Ito, R. Observation of the Stark effect in coupled quantum wells by electroluminescence and circularly polarized photoluminescence excitation spectroscopy. *J. Appl. Phys.* **75**, 7476–7481 (1994).
- Negoita, V., Snoke, D. W. & Eberl, K. Huge density-dependent blueshift of indirect excitons in biased coupled quantum wells. *Phys. Rev. B* **61**, 2779–2783 (2000).
- Moskalenko, S. A. & Snoke, D. W. *Bose-Einstein Condensation of Excitons and Biexcitons* (Cambridge Univ. Press, Cambridge, 2000).
- Bagaev, V. S., Keldysh, L. V., Sibeldin, N. N. & Tsvetkov, V. A. Phonon wind drag of excitons and electron-hole drops. *Sov. Phys. JETP* **70**, 702–716 (1976).
- Hensel, J. C. & Dynes, R. C. Interaction of electron-hole drops with ballistic phonons in heat pulses: The phonon wind. *Phys. Rev. Lett.* **39**, 969–972 (1977).
- Greenstein, M. & Wolfe, J. P. Anisotropy in the shape of the electron-hole-droplet cloud in germanium. *Phys. Rev. Lett.* **41**, 715–719 (1978).
- Snoke, D. W., Rühle, W. W., Köhler, K. & Ploog, K. Spin flip of excitons in GaAs quantum wells. *Phys. Rev. B* **55**, 13789–13794 (1997).

Acknowledgements

This work has been supported by the National Science Foundation and by the Department of Energy. We thank V. Negoita for early contributions to these experiments.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.S. (e-mail: snoke@pitt.edu).

The origin of the anomalous superconducting properties of MgB₂

Hyoungh Joon Choi*, David Roundy*†, Hong Sun*, Marvin L. Cohen*† & Steven G. Louie*†

* Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA

† Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Magnesium diboride¹ differs from ordinary metallic superconductors in several important ways, including the failure of conventional models² to predict accurately its unusually high transition temperature, the effects of isotope substitution on the critical transition temperature, and its anomalous specific heat^{3–5}. A detailed examination of the energy associated with the formation of charge-carrying pairs, referred to as the ‘superconducting energy gap’, should clarify why MgB₂ is different. Some early experimental studies have indicated that MgB₂ has multiple gaps^{3–9}, but past theoretical studies^{10–16} have not explained from first principles the origin of these gaps and their effects. Here we report an *ab initio* calculation of the superconducting gaps in MgB₂ and their effects on measurable quantities. An important feature is that the electronic states dominated by orbitals in the boron plane couple strongly to specific phonon modes, making pair formation favourable. This explains the high transition temperature, the anomalous structure in the specific heat, and the existence of multiple gaps in this material. Our analysis suggests comparable or higher transition temperatures may result in layered materials based on B, C and N with partially filled planar orbitals.

MgB₂ is a metal with a layer structure (Fig. 1a). The boron atoms form honeycombed layers, and the magnesium atoms are located above the centre of the hexagons in-between the boron planes. The electronic states at the Fermi level, which are the highest occupied electronic states, are mainly either σ - or π -bonding boron orbitals (Fig. 1b–d). The σ -bonding states are confined in the boron planes.

Thus, MgB₂ may be unique with partially occupied σ -bonding states in a layer structure. Because the charge distribution of the σ -bonding states is not symmetrical with respect to the in-plane positions of boron atoms, the σ -bonding states couple very strongly to the in-plane vibration of boron atoms (Fig. 1e). We show that this strong coupling results in strong electron-pair formation of the σ -bonding states with an average energy gap Δ of 6.8 meV. This strong pairing, which is confined to the boron planes and to only parts of the Fermi surface, is the principal contribution responsible for the superconductivity. However, the π -bonding states on the remaining parts of the Fermi surface form much weaker pairs with an average Δ of 1.8 meV. This pairing is enhanced by the coupling to the σ -bonding states.

Our present theoretical work is based on the strong coupling formalism of superconductivity established by Eliashberg^{17–19}. The Eliashberg formalism¹⁷ is a more general extension of the original formulation of the Bardeen–Cooper–Schrieffer (BCS) theory, which is based on the mechanism for pairing that involves an attractive interaction between electrons mediated by lattice vibrations. This approach is able to reproduce successfully the superconducting transition temperature of MgB₂ after detailed material properties are used¹⁶. This particular study¹⁶ provides only the transition temperature T_c , and does not give information on the superconducting energy gap (which is zero at T_c) or the temperature dependence of any measured quantities. However, the same general approach formally provides us with the nonlinear equations^{18,19} for the superconducting energy gap at temperatures below the transition temperature. We solved this set of equations using an iterative technique²⁰ to obtain the full crystal momentum $\mathbf{k}$ and temperature dependence of the energy gap of MgB₂ from first principles. A two-gap structure is assumed in model calculations¹⁴, and the ratio of the gap sizes obtained is consistent with the results presented here.

We have evaluated the properties of the superconducting energy gap of MgB₂ on the Fermi surface at low temperature (Fig. 2a, b). In MgB₂, the Fermi surface consists of two sheets from the σ -bonding states of boron $p_{x,y}$ orbitals, and two sheets from the π -bonding states of boron p_z orbitals. The calculation of the energy gap is made without any assumption of its functional shape on the Fermi surface. The resulting superconducting energy gap has *s*-wave

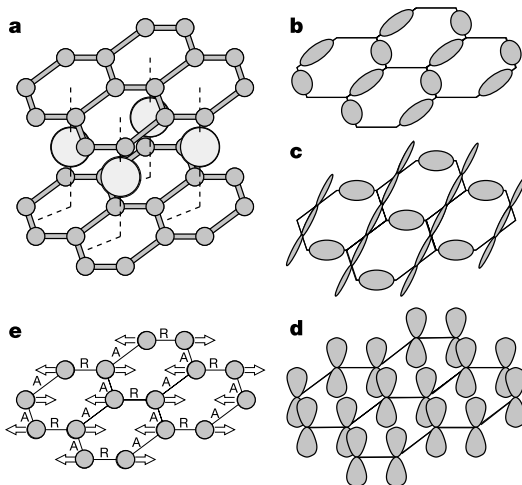


Figure 1 Crystal structure of MgB₂, electronic states at the Fermi level, and a vibrational mode of boron atoms. **a**, Crystal structure of MgB₂. Boron atoms form honeycomb planes, and magnesium atoms occupy the centres of the hexagons in-between boron planes. **b, c**, σ -bonding states at the Fermi level derived from boron $p_{x,y}$ orbitals. **d**, A π -bonding state at the Fermi level derived from boron p_z orbitals. **e**, A vibrational mode of boron atoms that couples strongly to σ -bonding electronic states at the Fermi level. As boron

atoms move in the arrow directions, shortened bonds, marked with ‘A’, become attractive to electrons, whereas elongated bonds, marked with ‘R’, become repulsive. The σ -bonding states (**b, c**) couple strongly to the vibrational mode because they are mainly located in either the attractive or the repulsive bondings of the mode. The π -bonding states (**d**) do not couple strongly to this mode.

symmetry (that is, the gap is of the same sign and non-zero everywhere on the Fermi surface), but the size of the gap changes greatly on the different sections of the Fermi surface. The magnitude of the energy gap at 4 K ranges from 6.4 to 7.2 meV on the σ sheets, and from 1.2 to 3.7 meV on the π sheets (Fig. 2a, b). The average values of the gap are 6.8 meV for the σ sheets and 1.8 meV for the π sheets. In experimental measurements, there has been a debate on the number of gaps^{3–9,21–26}. Our result is consistent with the recent experiments reporting two gaps, ranging from 1.5 to 3.5 meV for the small gap and 5.5 to 8 meV for the large gap^{3–9,26}.

The variation of the superconducting energy gap on the Fermi surface can be measured by techniques such as high-resolution angle-resolved photoelectron spectroscopy. Moreover, as the σ -bonding states are confined to the boron planes, the strong pairing gap of around 6.8 meV is associated with these planes (Fig. 2c). In Fig. 2c, we introduce the concept of a local gap distribution $\rho(\mathbf{r}, \Delta)$ at position $\mathbf{r}$ given by $\rho(\mathbf{r}, \Delta) = \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r})|^2 \delta(\Delta - \Delta_{\mathbf{k}})$, where $\psi_{\mathbf{k}}(\mathbf{r})$ is the electron wavefunction with crystal momentum $\mathbf{k}$. Our result shows that the small gaps should be seen preferentially in tunnelling experiments along the c axis, as indicated in some recent measurements^{6,7}.

Figure 3a depicts the calculated superconducting energy gaps at various temperatures below the transition temperature. The energy gap of the σ -bonding states and that of the π -bonding states show different temperature dependences. Compared to the small energy gap of the π -bonding states, the large energy gap of the σ -bonding states changes more slowly at low temperature, but more rapidly near the transition temperature. Both the π and σ gaps vanish at the same transition temperature, although their values are greatly different at low temperatures²⁷. This temperature dependence of the superconducting energy gaps explains recent tunnelling, optical and specific-heat measurements^{3–9}.

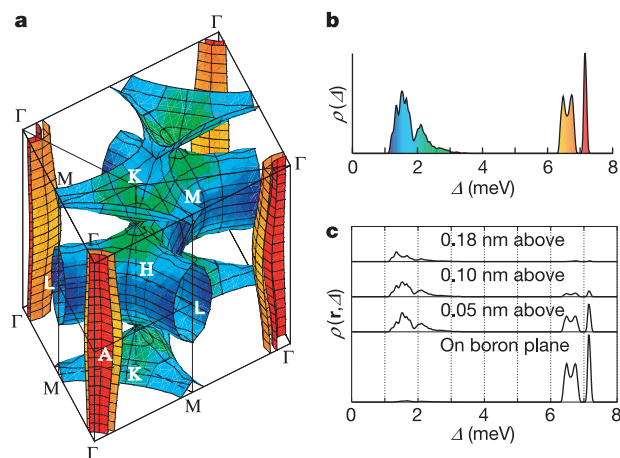


Figure 2 The superconducting energy gap of MgB_2 . **a**, **b**, The superconducting energy gap on the Fermi surface at 4 K given using a colour scale (**a**), and the distribution of gap values at 4 K (**b**). The Fermi surface of MgB_2 consists of four distinctive sheets. Two σ sheets ('cylinders'), derived from the σ -bonding p_{xy} orbitals of boron, are shown split into eight pieces around the four vertical Γ – Γ lines. Two π sheets ('webbed tunnels'), derived from the π -bonding p_z orbitals of boron, are shown around K–M and H–L lines (upper and lower K–M lines are equivalent). The superconducting energy gap is ~ 7.2 meV on the narrower σ cylindrical sheet, shown in red, with variations of less than 0.1 meV. On the wider σ cylindrical sheet, shown in orange, the energy gap ranges from 6.4 to 6.8 meV, having a maximum near Γ and a minimum near A. On the π sheets, shown in green and blue, the energy gap ranges from 1.2 to 3.7 meV. The density of states at the Fermi energy is 0.12 states per (eV atom spin), 44% of which comes from the σ sheets and the other 56% comes from the π sheets. **c**, Local distribution of the superconducting energy gap on a boron plane and on planes at 0.05, 0.10 and 0.18 nm above a boron plane, respectively.

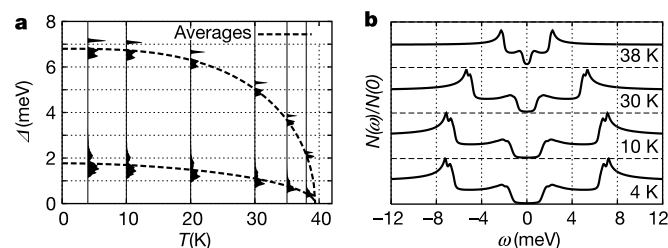


Figure 3 Calculated temperature dependence of the superconducting gaps and the quasiparticle density of states. **a**, Temperature dependence of the superconducting gaps. Vertical solid curves represent the distribution of the superconducting gap values at various temperatures from 4 K to 38 K. Dashed curves are of the form $\Delta(T) = \Delta(0) \times (1 - (T/T_c)^p)^{1/2}$ fitted separately to the calculated average energy gap of the σ -bonding states and that of the π -bonding states. For the σ sheets, $\Delta(0) = 6.8$ meV ($2\Delta(0)/k_B T_c = 4.0$) (k_B = Boltzmann's constant) and $p = 2.9$. For the π sheets, $\Delta(0) = 1.8$ meV ($2\Delta(0)/k_B T_c = 1.06$) and $p = 1.8$. **b**, The quasiparticle density of states at various temperatures. The quasiparticle density of states $N(\omega)$ for the quasiparticle energy ω is given by $N(\omega)/N(0) = \text{Re}[(\omega + i\Gamma)/(\omega + i\Gamma)^2 - \Delta(\mathbf{k}, \omega)^2]^{1/2}$, where $N(0)$ is the electron density of states at the Fermi level, $i = (-1)^{1/2}$, and $\langle \dots \rangle$ indicates an average over a surface of constant ω . This curve is obtained from the calculated gap function $\Delta(\mathbf{k}, \omega)$ and an assumed finite lifetime Γ of 0.1 meV.

The superconducting energy gap determines the quasiparticle density of states. The quasiparticle energy is the excitation energy of a system when an electron is added or removed. In a superconductor, the quasiparticle energy is equal to, or greater than, the superconducting energy gap Δ . Because the energy gap differs considerably for the σ - and π -bonding states in MgB_2 , the density of quasiparticle excitations as a function of energy shows two thresholds (Fig. 3b). Only π -bonding quasiparticle states are allowed for energies between the minimal superconducting energy gap of the π -bonding states and that of the σ -bonding states. For energies above the minimal superconducting energy gap of the σ -bonding states, quasiparticle excitation becomes possible for both the σ - and π -bonding states. The quasiparticle density of states can be deduced experimentally from tunnelling experiments and var-

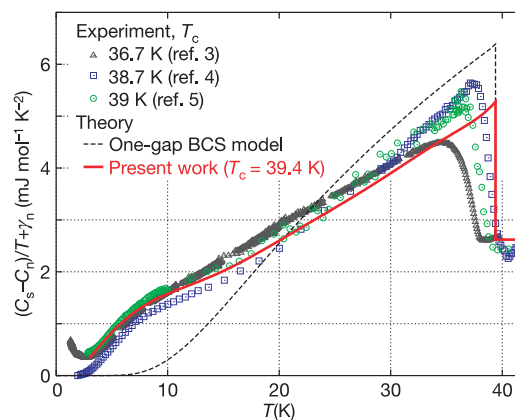


Figure 4 The specific heat of MgB_2 . The measured and calculated electronic contribution to the specific heat divided by temperature are plotted as a function of temperature. The red solid curve represents the result of our calculation. The specific heat difference $(C_S - C_N)$ between the superconducting and normal states is obtained by $C_S - C_N = -T(d^2/dT^2)(F_S - F_N)$ from the corresponding free energy difference $(F_S - F_N)$ which is calculated using a generalized Bardeen–Stephen formula²⁸. The normal-state specific heat is calculated to be $C_N = Y_N T$ with $Y_N = 2.62 \text{ mJ mol}^{-1} \text{ K}^{-2}$ (ref. 16). Symbols are the results of experimental measurements^{3–5}, and the dashed curve is the standard one-gap BCS prediction corresponding to a transition temperature of 39.4 K.

ious spectroscopic measurements^{6–9}, but a direct quantitative comparison requires knowledge of various physical parameters involved in a specific experiment.

A direct quantitative comparison between theory and measurement is possible, however, for the specific heat. The measured specific heat^{3–5} of MgB₂ at low temperature is substantial, and a large hump appears at about 10 K that is inconsistent with a one-gap BCS model. We have calculated the specific heat from the free energy of the superconducting state. The overall shape and magnitude of our calculated specific-heat curve agrees very well with the experimental data, especially below 30 K (Fig. 4). We find that the low-temperature hump in our calculated curve and in the experimental data is caused by the existence of low-energy excitations across the small superconducting energy gap of the π -bonding states. □

Received 14 November 2001; accepted 31 May 2002; doi:10.1038/nature00898.

- Nagatsuki, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. Superconductivity at 39 K in magnesium diboride. *Nature* **410**, 63–64 (2001).
- Hinks, D. G., Claus, H. & Jorgensen, J. D. The complex nature of superconductivity in MgB₂ as revealed by the reduced total isotope effect. *Nature* **411**, 457–460 (2001).
- Wang, Y., Plackowski, T. & Junod, A. Specific heat in the superconducting and normal state (2–300 K, 0–16 T), and magnetic susceptibility of the 38 K superconductor MgB₂. *Physica C* **355**, 179–193 (2001).
- Bouquet, F., Fisher, R. A., Phillips, N. E., Hinks, D. G. & Jorgensen, J. D. Specific heat of Mg¹¹B₂: evidence for a second energy gap. *Phys. Rev. Lett.* **87**, 047001–1–047001–4 (2001).
- Yang, H. D. *et al.* Order parameter of MgB₂: a fully gapped superconductor. *Phys. Rev. Lett.* **87**, 167003–1–167003–4 (2001).
- Szabo, P. *et al.* Evidence for two superconducting energy gaps in MgB₂ by point-contact spectroscopy. *Phys. Rev. Lett.* **87**, 137005–1–137005–4 (2001).
- Giubileo, F. *et al.* Two-gap state density in MgB₂: a true bulk property or a proximity effect? *Phys. Rev. Lett.* **87**, 177008–1–177008–4 (2001).
- Chen, X. K., Konstantinovi, M. J., Irwin, J. C., Lawrie, D. D. & Franck, J. P. Evidence for two superconducting gaps in MgB₂. *Phys. Rev. Lett.* **87**, 157002–1–157002–4 (2001).
- Tsuda, S. *et al.* Evidence for a multiple superconducting gap in MgB₂ from high-resolution photoemission spectroscopy. *Phys. Rev. Lett.* **87**, 177006–1–177006–4 (2001).
- Kortus, J., Mazin, I. I., Belashchenko, K. D., Antropov, V. P. & Boyer, L. L. Superconductivity of metallic boron in MgB₂. *Phys. Rev. Lett.* **86**, 4656–4659 (2001).
- An, J. M. & Pickett, W. E. Superconductivity of MgB₂: covalent bonds driven metallic. *Phys. Rev. Lett.* **86**, 4366–4369 (2001).
- Bohnen, K.-P., Heid, R. & Renker, B. Phonon dispersion and electron-phonon coupling in MgB₂ and AlB₂. *Phys. Rev. Lett.* **86**, 5771–5774 (2001).
- Yildirim, T. *et al.* Giant anharmonicity and nonlinear electron-phonon coupling in MgB₂: a combined first-principles calculation and neutron scattering study. *Phys. Rev. Lett.* **87**, 037001–1–037001–4 (2001).
- Liu, A. Y., Mazin, I. I. & Kortus, J. Beyond Eliashberg superconductivity in MgB₂: anharmonicity, two-phonon scattering, and multiple gaps. *Phys. Rev. Lett.* **87**, 087005–1–087005–4 (2001).
- Kong, Y., Dolgov, O. V., Jepsen, O. & Andersen, O. K. Electron-phonon interaction in the normal and superconducting states of MgB₂. *Phys. Rev. B* **64**, 020501–1–020501–4 (2001).
- Choi, H. J., Roundy, D., Sun, H., Cohen, M. L. & Louie, S. G. First-principles calculation of the superconducting transition in MgB₂ within the anisotropic Eliashberg formalism. *Phys. Rev. B* **66**, 020513–1–020513–4 (2002).
- Eliashberg, G. M. Interactions between electrons and lattice vibrations in a superconductor. *Zh. Eksp. Teor. Fiz.* **38**, 966–976 (1960); *Sov. Phys. JETP* **11**, 696–702 (1960).
- Allen, P. B. & Mitrović, B. In *Solid State Physics* (eds Ehrenreich, H., Seitz, F. & Turnbull, D.) Vol. 37 1–92 (Academic, New York, 1982).
- Carbotte, J. P. Properties of boson-exchange superconductors. *Rev. Mod. Phys.* **62**, 1027–1157 (1990).
- Marsiglio, F., Schossmann, M. & Carbotte, J. P. Iterative analytic continuation of the electron self-energy to the real axis. *Phys. Rev. B* **37**, 4965–4969 (1988).
- Karapetrov, G., Iavarone, M., Kwok, W. K., Crabtree, G. W. & Hinks, D. G. Scanning tunneling spectroscopy in MgB₂. *Phys. Rev. Lett.* **86**, 4374–4377 (2001).
- Sharoni, A., Felner, I. & Millo, O. Tunneling spectroscopy and magnetization measurements of the superconducting properties of MgB₂. *Phys. Rev. B* **63**, 220508–1–220508–4 (2001).
- Rubio-Bollinger, G., Suderow, H. & Vieira, S. Tunneling spectroscopy in small grains of superconducting MgB₂. *Phys. Rev. Lett.* **86**, 5582–5584 (2001).
- Schmidt, H., Zasadzinski, J. F., Gray, K. E. & Hinks, D. G. Energy gap from tunneling and metallic contacts onto MgB₂: possible evidence for a weakened surface layer. *Phys. Rev. B* **63**, 220504–1–220504–4 (2001).
- Takahashi, T., Sato, T., Souma, S., Muranaka, T. & Akimitsu, J. High-resolution photoemission study of MgB₂. *Phys. Rev. Lett.* **86**, 4915–4917 (2001).
- Buza, C. & Yamashita, T. Review of superconducting properties of MgB₂. *Supercond. Sci. Technol.* **14**, R115–R146 (2001).
- Suhl, H., Matthias, B. T. & Walker, L. R. Bardeen-Cooper-Schrieffer theory of superconductivity in the case of overlapping bands. *Phys. Rev. Lett.* **3**, 552–554 (1959).
- Bardeen, J. & Stephen, M. Free-energy difference between normal and superconducting states. *Phys. Rev. A* **136**, 1485–1487 (1964).

Acknowledgements

This work was supported by the National Science Foundation and by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy.

Computational resources have been provided by the National Science Foundation at the National Center for Supercomputing Applications and by the National Energy Research Scientific Computing Center. H. S. acknowledges financial support from the Berkeley Scholar Program funded by the Tang Family Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.G.L. (e-mail: sglozie@uclink.berkeley.edu).

The influence of a chemical boundary layer on the fixity, spacing and lifetime of mantle plumes

A. Mark Jellinek & Michael Manga

Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA

Seismological observations provide evidence that the lowermost mantle contains superposed thermal and compositional boundary layers¹ that are laterally heterogeneous^{2,3}. Whereas the thermal boundary layer forms as a consequence of the heat flux from the Earth's outer core, the origin of an (intrinsically dense) chemical boundary layer remains uncertain⁴. Observed zones of 'ultra-low' seismic velocity⁵ suggest that this dense layer may contain metals^{6,7} or partial melt⁸, and thus it is reasonable to expect the dense layer to have a relatively low viscosity. Also, it is thought that instabilities in the thermal boundary layer could lead to the intermittent formation and rise of mantle plumes. Flow into ascending plumes can deform the dense layer, leading, in turn, to its gradual entrainment^{9–14}. Here we use analogue experiments to show that the presence of a dense layer at the bottom of the mantle induces lateral variations in temperature and viscosity that, in turn, determine the location and dynamics of mantle plumes. A dense layer causes mantle plumes to become spatially fixed, and the entrainment of low-viscosity fluid enables plumes to persist within the Earth for hundreds of millions of years.

Convective motions driven by core cooling have a structure that is three-dimensional and time-dependent. Consequently, the dynamics of the interaction between this flow and an underlying dense layer is complex. Because of the computational challenge of resolving small (kilometre) length scales while tracking viscosity and density interfaces¹⁵, numerical simulations for conditions appropriate to mantle convection are typically limited to two dimensions^{12,13}, though three-dimensional simulations are currently being performed¹⁴. Laboratory experiments are thus often used to study thermochemical convection, and the experiments presented here extend previous investigations^{9,10,16} to the situation in which the dense layer is thin and has a low viscosity. Here we need to distinguish 'plumes' from 'thermals': we use 'plume' to describe buoyant upwellings (or downwellings) that extend continuously from the hot (or cold) boundary layer, and 'thermal' to indicate a discrete buoyant blob. Under conditions of thermal equilibrium, more-viscous cold fluid descends in narrow sheet-like plumes¹⁷, whereas lower-viscosity hot fluid ascends mostly in thermals^{18,19}.

Our experiments are performed in the tank illustrated in Fig. 1 (the legend of Fig. 1 also defines the parameters and variables). An initial series of 23 experiments without a dense layer is performed to provide a framework to interpret the more complicated situation

including a dense layer. Both free-slip and no-slip bottom boundaries are used. The tank is filled with a single layer of polybutene oil (Chevron Oronite 16500), which is a very viscous, colourless newtonian fluid. Before the start of each experiment the oil layer is heated from below and cooled from above until thermal equilibrium is achieved. Time-lapse video, and measurements of the surface heat flux, surface temperature and interior temperature as a function of time are used to characterize flows quantitatively. Thermal equilibrium is indicated by constant and equal heat fluxes at the roof and floor of the tank, and by an interior fluid temperature that is statistically constant over time. Experiments are run for several days, corresponding to thousands of plume rise times. Here, a 'rise time' is defined as the timescale for a thermal or plume to ascend from the hot thermal boundary layer to the roof of the tank.

The second series of experiments is designed specifically to understand the interaction between convection driven by heat flow from the core and an underlying dense, low-viscosity layer under conditions appropriate to the Earth's mantle. We first allow the convecting fluid to reach thermal equilibrium, and then introduce dense, low-viscosity, miscible soybean oil through a 4-mm-diameter port in a corner of the tank (Fig. 2). This dense fluid (coloured red) spreads over the floor of the tank, driving no apparent large-scale circulation, and forms a nearly isothermal layer that is 0.5 cm high. Following its introduction, this 'dense layer' is deformed by flow of thermal boundary layer fluid into ascending thermals of polybutene oil (the 'ambient fluid'). Figure 2 shows circular embayments separated by sharp 'divides' that formed early in the evolution of the red, dense layer. The dynamic coupling between topography and motions driven by lateral temperature variations stabilize the pattern of flow over timescales much longer than a plume rise time; buoyant fluid is able to ascend along the sloping interface with the low-viscosity soybean oil more easily than rise vertically into the more-viscous ambient fluid. To confirm this mechanism and to understand the influences of topography on the dense layer and the viscosity ratio λ_d , we performed 18 additional experiments. Figure 3 shows that smaller (but still finite) topographic relief is needed to stabilize the positions of plumes when λ_d

is large. The spacing between the centres of the embayments remains approximately constant, apparently governed by the wavelength of the first Rayleigh–Taylor instability of the thermal boundary layer (discussed in more detail later). Stable flows are enhanced because entrainment of the dense layer establishes narrow, cylindrical low-viscosity conduits beneath ascending thermals (Fig. 4), causing a transition from an unsteady flow dominated by upwelling thermals to a steady flow governed by axisymmetric plumes with persistent trailing conduits. Finite deformation of the dense layer is needed to stabilize plume locations, because a flat boundary permits only thermals to form (Fig. 4b).

Our results suggest that the persistence, fixity and composition of mantle plumes can depend on the mechanics of entrainment from, and the longevity of, a dense layer. Accordingly, we extend previous studies^{9,11,16,20,21} by developing scaling relationships for the radius of entrained low-viscosity tendrils and the erosion rate of the dense layer. We first consider entrainment by flow into thermals forming at the hot boundary. Second, we address entrainment by flow into established conduits.

Flow into an ascending thermal or an established plume conduit drags a layer of material from the dense layer. Assuming $\Delta\rho > 0$, $\lambda_d \gg 1$ and that the entrained tendril of fluid is thin relative to the thickness of the thermal boundary layer, a balance of viscous (drag)

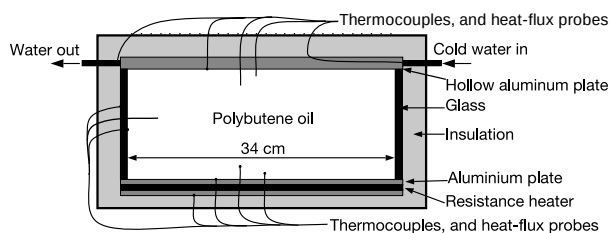


Figure 1 A schematic illustration of the apparatus used for the experiments shown in Figs 2 and 4 (ref. 26). The interior of the tank has a depth of 17 cm. Five dimensionless parameters are used to characterize our experiments and scale the results to the Earth's mantle. (1) The Rayleigh number, $Ra = g\rho\alpha\Delta TH^3/\mu_c\kappa > 10^5$; (2) the Prandtl number, $Pr = \mu_c/\rho\kappa > 10^4$. Consequently, flows are laminar (low Reynolds number), time-dependent and composed predominantly of blobs ascending (or descending) from the hot (or cold) boundaries²⁷. Here, H is the combined height of both fluid layers, g the acceleration due to gravity, ΔT the temperature difference across the height H , ρ the fluid density, α the coefficient of thermal expansion, κ the thermal diffusivity, and μ_c , μ_h and μ_d are the viscosities evaluated in the ambient fluid, at the interface with the hot boundary and in the dense layer, respectively. (3) The ratio of the buoyancies of the dense layer to the thermal boundary layer, $B = \Delta\rho/\rho\alpha\Delta T$, where $\Delta\rho$ is the intrinsic density difference between the dense layer and overlying fluid, indicates the gravitational stability of the dense layer. $B < 1$ can still be stable because ΔT is based on the full temperature difference across the convecting system rather than the smaller temperature difference across the hot thermal boundary layer¹⁶. We choose this definition because ΔT is an externally controlled variable. (4) The viscosity ratio $\lambda_d = \mu_c/\mu_d$, and (5) the viscosity ratio $\lambda_h = \mu_c/\mu_h$; both of these ratios depend on temperature.

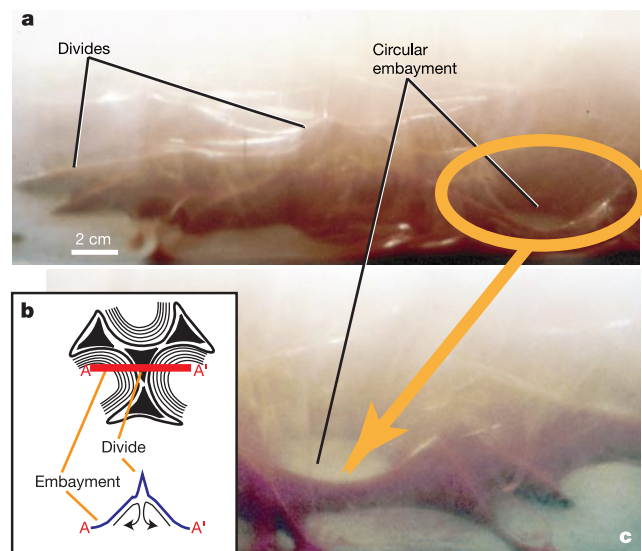


Figure 2 Photographs showing an oblique view of the temporal evolution of a dense, low-viscosity layer of soybean oil (coloured red) emplaced at the base of a vigorously convecting layer of polybutene oil (colourless). Here, $Ra = 3.6 \times 10^7$, $B = 0.56$, $\lambda_d = 1.1 \times 10^2$ and $\lambda_h = 2.7$. The ratio of the height of the dense layer D (0.5 cm) to the total height of the system H (17 cm) is chosen to be an analogue for the ratio of the height of dense layer (a few hundred kilometres, say) to the height of the mantle (3,000 km). Initially, the spreading of dense soybean oil imparts a thermal transient to the bottom boundary. However, because the height of the dense layer is comparable to the thickness of the thermal boundary layer, thermal equilibrium is reacquired over the timescale for emplacement of the layer. **a**, Subsequent entrainment from the dense layer by flow into ascending thermals produces circular embayments and divides that, in turn, determine the style of convection from the hot boundary. Dynamically, horizontal temperature gradients drive flow from the tank floor, up the sides of embayments to peaks on the divides, where upwellings into the ambient fluid become fixed (compare Fig. 3). **b**, A schematic contour map and cross-section illustrating the topography in **a**. Arrows indicate motions within the dense layer, and show the nature of the predominantly mechanical coupling (but also thermal coupling—see Fig. 4) between the two layers. **c**, Approximately 120 plume rise times later in the experiment (equivalent to 3.6 Gyr of Earth's history), upwellings remain controlled by the positions of divides and embayments formed at the start of the experiment (although embayments have been eroded significantly).

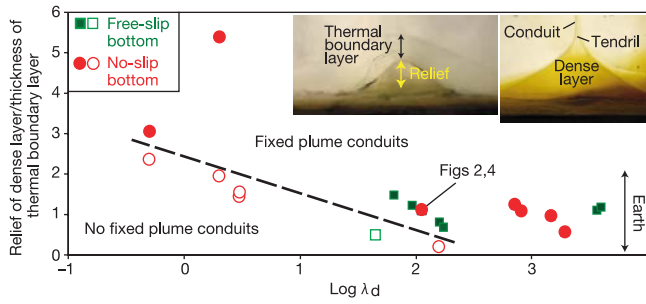


Figure 3 Plot summarizing the conditions leading either to spatially fixed (filled symbols) or not-fixed (open symbols) plumes. Except where indicated, the data are obtained using the method and apparatus of ref. 16. Circles, from experiments with a no-slip bottom boundary condition; squares, from experiments in which a free-slip lower boundary is obtained using the method of ref. 28. When λ_d is large ($>10^2$), buoyant fluid ascends along the sloping interface with the dense layer (rather than vertically)²⁹ because the underlying low-viscosity soybean oil presents a smaller retarding drag force than the more-viscous ambient fluid. When λ_d is small (~ 1), the viscous stresses due to the dense layer and interior fluid are similar. So for $\lambda_d < 10$, the relief must be comparable to, or greater than, the thickness of the thermal boundary layer to cause the convection to be governed by steady flow into long-lived, stationary plume conduits. Smaller topographic relief results in unsteady flow, in which plumes have a spacing and lifetime that is approximately stochastic. As λ_d becomes very large, however, increasingly less (but still finite; Fig. 4) relief is required to cause the convection to become governed by fixed plumes. The dashed line schematically separates conditions leading to fixed or not-fixed plumes. This line cannot be extrapolated to the horizontal axis because topographic relief on the dense layer is required for fixed plumes.

and buoyancy forces¹¹ leads to a scale for the radius of an entrained tendrils:

$$l \approx \left(\frac{\mu_d V}{g \Delta \rho} \right)^{1/2} = \left(\frac{H^3 V}{\kappa B R a \lambda_d} \right)^{1/2} \quad (1)$$

For thermals, an appropriate velocity, V , is the rise speed $V_{th} \approx \kappa H^{-1} Ra^{1/3}$ (ref. 22) and equation (1) thus becomes:

$$l_{th} \approx \left(\frac{H^2}{\lambda_d B R a^{2/3}} \right)^{1/2} \quad (2)$$

An estimate of the average erosion rate of the dense layer by flow into rising thermals (or conduits) depends on the area of the region of thermal-boundary-layer fluid feeding each type of upwelling. Hence, we require constraints on the horizontal spacing between upwellings. The measured average horizontal spacing between 6–8 established conduits in the experiment shown in Figs 2 and 4 is approximately constant, and between 8 and 10 cm. For a critically unstable thermal boundary layer of thickness $10 H R a^{-1/3}$ (ref. 22) embedded between an underlying low-viscosity dense layer and overlying more-viscous ambient fluid, linear stability theory²³ constrains the most unstable wavelength:

$$L \approx \frac{20\pi H}{C} \left(\frac{\lambda_h}{R a} \right)^{1/3} \quad (3)$$

where the constant $C \approx 0.6$ depends on λ_h , λ_d and B ; here, L is ≈ 7.8 cm. The consistency between measurements and theory suggests that the spacing between plumes is governed by the first Rayleigh–Taylor instability of the buoyant thermal boundary layer following the introduction of the dense layer. The fact that conduit spacing can be predicted with linear theory is important for two reasons. First, this finding supports our argument that the dynamic coupling between rising thermals and an underlying dense layer leads to upwellings that remain fixed spatially. Second, a scale for the horizontal spacing between upwellings may now be written explicitly in terms of parameters controlled externally. Consequently, the

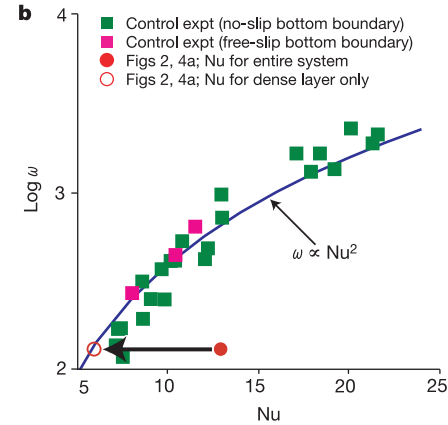
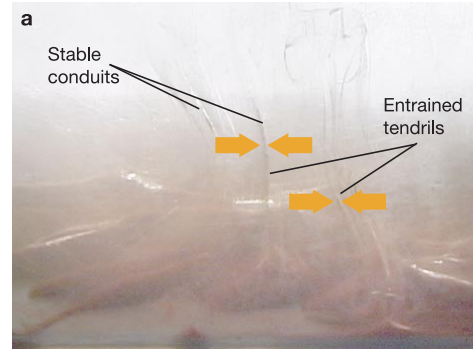


Figure 4 Illustration of the spatial and temporal structure of plumes. **a**, Photograph illustrating persistent, axisymmetric, low-viscosity conduits behind rising thermals. Once formed, conduits govern the flow of buoyant fluid from the hot boundary and persist until the dense layer is eroded completely. Hence, the conduits are regarded as stable structures. The arrows indicate entrained tendrils of dense fluid. Tendril radii, which are measured near the bases of 6–8 established conduits during an experiment, are in the range 0.25–1 mm with a mean of 0.6 mm, and constrain the scaling constant, $\Lambda = 0.7$. The average vertical flux per area of dense fluid, which is obtained from time-lapse video, is about 3–5 mm h⁻¹. **b**, The frequency of thermal formation, ω , made dimensionless with the conductive timescale, H^2/κ , as a function of Nusselt number, Nu (the measured dimensionless heat flux). ω is obtained with heat-flux probes, thermocouples and visually. Solid squares are control experiments performed without a dense layer with no-slip and free-slip lower boundaries. The theoretical curve ($\omega \propto Nu^2$) is in agreement with the control experiments and is compatible with convection from the hot boundary being in the form of intermittent thermals¹⁸. The data shown are from the dense-layer experiment shown in **a** and in Fig. 2. Filled and open circles are based on Nu for the entire system and for the dense layer, respectively. At the start of the experiment (Fig. 2a) ω is reduced by about a factor of 5, compared with a flow transporting the same amount of heat without a dense layer. The measured frequency implies that thermals in the ambient fluid are coupled to the flow in the dense layer. Once conduits become established, however, no characteristic frequency is observed, and the flow is steady. After the erosion of the dense layer, the flow evolves back to intermittent thermals as in the control experiments.

erosion rate of dense layer by flow into rising thermals is

$$q_{th} \approx V_{th} \frac{l^2}{L^2} = \frac{C^2}{20^2 \pi^2} \left(\frac{\kappa}{H} \right) \frac{Ra^{1/3}}{\lambda_d B R a^{2/3}} \quad (4)$$

When convection is governed by established plume conduits, scaling arguments for both the tendrils radius l_{pl} and the erosion rate q_{pl} depend on the spacing between plumes^{20,21}. In this case, the flow velocity²¹ V_{pl} is given approximately by $(\kappa/H)(L/H)Ra^{2/3}$. Thus, from equations (1) and (3) the average tendrils radius is:

$$l_{pl} = 10.2 \Lambda H \left(\frac{\lambda_h^{1/3}}{\lambda_d B R a^{2/3}} \right)^{1/2} \quad (5)$$

The measured l_{pl} is 0.6 mm (Fig. 4 legend), implying that the constant $\Lambda = 0.7$.

Time-lapse video is used to measure the average erosion rate of the dense layer by flow into established conduits. The time-derivative of the position of the density interface indicates an average erosion rate of 3–5 mm h⁻¹. From equations (3) and (5), the erosion rate is

$$q_{pl} = V_{pl} \frac{l_{pl}^2}{L^2} = \frac{\Lambda^2 \kappa Ra^{1/3}}{H \lambda_d B} \quad (6)$$

which is consistent with refs 9 and 16, and which predicts $q_{pl} = 5 \text{ mm h}^{-1}$ (Fig. 3).

Applying our results to the Earth requires constraints on Ra , B , λ_d and λ_h . Ra for mantle convection driven by core cooling is in the range 10^7 – 10^8 . The data in Fig. 3 imply that λ_d needs to be in excess of 10^1 – 10^2 in order for fixed plumes to occur. In addition, we take $\lambda_h = 10^2$ (ref. 24) and $B = 0.5$, along with typical values of $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ and $H = 3 \times 10^6 \text{ m}$. Assuming that mantle starting plumes are dynamically similar to thermals, equations (2) and (4) imply that tendrils entrained by flow into starting plumes are less than 1 km in radius, and that the associated erosion rates are of the order of 10^{-8} – $10^{-7} \text{ km Myr}^{-1}$. On the other hand, because $V_{pl} \gg V_{th}$, equations (5) and (6) predict that tendrils entrained into established low-viscosity conduits are 5–10 km in radius, and the resultant entrainment rate is expected to be of the order of 10^{-2} – $10^{-3} \text{ km Myr}^{-1}$. If a typical dense layer is, say, 100–300 km high, these erosion rates imply that the layer could persist for times comparable to, or longer than, the age of the Earth. Finally, equation (3) predicts that, in the presence of a dense layer, the spacing between mantle plumes is of the order of 10^3 km . This value is compatible with the distribution of hotspots in the Pacific, which appear to overlie the ultra-low-velocity zone³, when their positions are extrapolated to the core–mantle boundary^{3,25,30}.

We have shown that a dense, low-viscosity compositional boundary layer at the base of the mantle may determine the location, relative fixity and spacing of mantle plumes. Moreover, we have argued that entrainment from this low-viscosity layer may influence the composition of plumes and be important in establishing long-lived, low-viscosity conduits extending from the thermal boundary layer at the base of the mantle to the base of the lithosphere. Our results may resolve a basic conundrum in mantle dynamics—that hotspots remain approximately fixed in space relative to each other for timescales much longer than the time required for a plume to rise through the mantle. □

Received 26 October 2001; accepted 15 July 2002; doi:10.1038/nature00979.

1. Lay, T., Williams, Q. & Garnero, E. J. The core–mantle boundary layer and deep Earth dynamics. *Nature* **392**, 461–468 (1998).
2. Wyssession, M. E. Large-scale structure at the core–mantle boundary from diffracted waves. *Nature* **382**, 244–248 (1996).
3. Castle, J. C., Creager, K. C., Winchester, J. P. & van der Hilst, R. D. Shear wave speeds at the base of the mantle. *J. Geophys. Res.* **105**, 21543–21557 (2000).
4. Gurnis, M., Wyssession, M. E., Knittle, E. & Buffett, B. A. (eds) *The Core–Mantle Region Monograph 28* (American Geophysical Union, Washington DC, 1998).
5. Garnero, E. J. Heterogeneity of the lowermost mantle. *Annu. Rev. Earth Planet. Sci.* **28**, 509–537 (2000).
6. Knittle, E. & Jeanloz, R. The Earth's core–mantle boundary: Results of experiments at high pressures and temperatures. *Science* **251**, 1438–1453 (1991).
7. Brandon, A. D. et al. Coupled ¹⁸⁶Os and ¹⁸⁷Os evidence for core–mantle interaction. *Science* **280**, 1570–1573 (1998).
8. Williams, Q. & Garnero, E. J. Seismic evidence for partial melt at the base of the Earth's mantle. *Science* **273**, 1528–1530 (1996).
9. Davaille, A. Simultaneous generation of hotspots and superswells by convection in a heterogeneous planetary mantle. *Nature* **402**, 756–760 (1999).
10. Olson, P. & Kincaid, C. Experiments on the interaction of thermal convection and compositional layering at the base of the mantle. *J. Geophys. Res.* **96**, 4347–4354 (1991).
11. Sleep, N. H. Gradual entrainment of a chemical layer at the base of the mantle by overlying convection. *Geophys. J. Int.* **95**, 437–447 (1998).
12. Montague, N. L. & Kellogg, L. H. Numerical models of a dense layer at the base of the mantle and implications for the geodynamics of D'. *J. Geophys. Res.* **105**, 11101–11114 (2000).
13. Farnetani, C. G. Excess temperature of mantle plumes: the role of chemical stratification across D'. *Geophys. Res. Lett.* **24**, 1583–1586 (1997).

14. Tackley, P. J. The strong heterogeneity caused by deep mantle layering. *Geochem. Geophys. Geosyst.* **3**, U1–U22 (2002).
15. van Keken, P. E. et al. A comparison of methods for the modeling of thermochemical convection. *J. Geophys. Res.* **102**, 22477–22495 (1997).
16. Gonnermann, H. M., Manga, M. & Jellinek, A. M. Dynamics and longevity of an initially stratified mantle. *Geophys. Res. Lett.* **29**, 10.1029/2002GL014851 (2002).
17. Neavel, K. E. & Johnson, A. M. Entrainment in compositionally buoyant plumes. *Tectonophysics* **200**, 1–15 (1991).
18. Manga, M. & Weeraratne, D. Experimental study of non-Boussinesq Rayleigh–Benard convection at high Rayleigh and Prandtl numbers. *Phys. Fluids* **11**, 2969–2976 (1999).
19. Sparrow, E. M., Husar, R. B. & Goldstein, R. J. Observations and other characteristics of thermals. *J. Fluid Mech.* **41**, 793–802 (1970).
20. Davaille, A. Two-layer thermal convection in miscible viscous fluids. *J. Fluid Mech.* **379**, 223–253 (1999).
21. Olson, P., Schubert, G. & Anderson, C. Structure of axisymmetric mantle plumes. *J. Geophys. Res.* **98**, 6829–6844 (1993).
22. Turner, J. S. *Buoyancy Effects in Fluids* (Cambridge Univ. Press, Cambridge, 1973).
23. Lister, J. R. & Kerr, R. C. The effect of geometry on the gravitational instability of a buoyant region of viscous fluid. *J. Fluid Mech.* **202**, 577–594 (1989).
24. Nataf, H.-C. Mantle convection, plates, and hotspots. *Tectonophysics* **187**, 361–377 (1991).
25. Sleep, N. H. Time dependence of mantle plumes: Some simple theory. *J. Geophys. Res.* **97**, 20007–20019 (1992).
26. Schaeffer, N. & Manga, M. Interaction of rising and sinking mantle plumes. *Geophys. Res. Lett.* **28**, 455–458 (2001).
27. Neavel, K. E. & Johnson, A. M. Entrainment in compositionally buoyant plumes. *Tectonophysics* **200**, 1–15 (1991).
28. Jellinek, A. M., Lenardic, A. & Manga, M. The influence of interior mantle temperature on the structure of plumes: Heads for Venus, Tails for the Earth. *Geophys. Res. Lett.* **29**, 10.1029/2001GL014624 (2002).
29. Martin, D. & Campbell, I. H. Laboratory modeling of convection in magma chambers: Crystallization against sloping floors. *J. Geophys. Res.* **93**, 7974–7988 (1988).
30. Steinberger, B. & O'Connell, R. J. Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution. *Geophys. J. Int.* **132**, 412–434 (1998).

Acknowledgements

We thank A. Davaille for comments on the previous version of the manuscript, and L. H. Kellogg, R. C. Kerr, M. A. Richards and N. Sleep for comments on earlier versions. This work was supported by the National Science Foundation and The Miller Institute for Basic Research in Science.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.M.J. (e-mail: markj@seismo.berkeley.edu).

Seismic evidence for catastrophic slab loss beneath Kamchatka

Vadim Levin*, Nikolai Shapiro*†, Jeffrey Park* & Michael Ritzwoller*†

* Department of Geology and Geophysics, Box 208109, Yale University, New Haven, Connecticut 06520, USA

† Department of Physics, University of Colorado, Boulder, Colorado 80309, USA

In the northwest Pacific Ocean, a sharp corner in the boundary between the Pacific plate and the North American plate joins a subduction zone running along the southern half of the Kamchatka peninsula¹ with a region of transcurrent motion along the western Aleutian arc^{1,2}. Here we present images of the seismic structure beneath the Aleutian–Kamchatka junction and the surrounding region, indicating that: the subducting Pacific lithosphere terminates at the Aleutian–Kamchatka junction; no relict slab underlies the extinct northern Kamchatka volcanic arc; and the upper mantle beneath northern Kamchatka has unusually slow shear wavespeeds. From the tectonic and volcanic evolution of Kamchatka over the past 10 Myr (refs 3–5) we infer that at least two episodes of catastrophic slab loss have occurred. About 5 to 10 Myr ago, catastrophic slab loss shut down island-arc volcanic

activity north of the Aleutian–Kamchatka junction. A later episode of slab loss, since about 2 Myr ago, seems to be related to the activity of the world's most productive island-arc volcano, Klyuchevskoy⁶. Removal of lithospheric mantle is commonly discussed in the context of a continental collision, but our findings imply that episodes of slab detachment and loss are also important agents in the evolution of oceanic convergent margins.

At convergent margins the descent of the lithospheric slab into the upper mantle may give way to tearing, detachment, and sinking of the slab material, accompanied by uplift, extension and distinctive magmatism^{7–9}. Ongoing slab tears are proposed for the Apennines⁸ and Taiwan⁹, and the foundering of lithospheric mantle has been documented beneath Tibet¹⁰ and the Alboran Sea¹¹. There is, therefore, growing evidence for lithospheric mantle detachment as a mechanism for the rejuvenation of continental lithosphere. Evidence for similar detachment of the mantle lithosphere in an oceanic margin setting is much weaker. Here, we present tomographic images of the shallow mantle from surface-wave dispersion data that, together with other geophysical and geochemical evidence, imply several episodes of slab removal beneath the Kamchatka peninsula of the Russian Far East. We argue that lithospheric detachment is essential to the evolution of this and other oceanic-margin subduction zones.

Deep subduction-related earthquakes range from 400 km beneath southernmost Kamchatka to ~200 km at the Aleutian–Kamchatka junction (AKJ), and slab dip decreases from 55° to 35° (ref. 12). A chain of volcanoes along the eastern coast of Kamchatka traces the 100-km depth-contour of the subducted Pacific lithosphere, culminating in a cluster of volcanoes inland of the AKJ. The distribution of volcanic rocks on the peninsula suggests geologically recent rearrangement of the associated subduction zones¹³. Before about 10 Myr ago, island-arc magmatism extended north of the AKJ along the mid-Kamchatka volcanic belt (Fig. 1). That volcanism changed character afterwards^{13,14} and is now extinct⁴.

The dynamics of the Kamchatka subduction zone reflect the

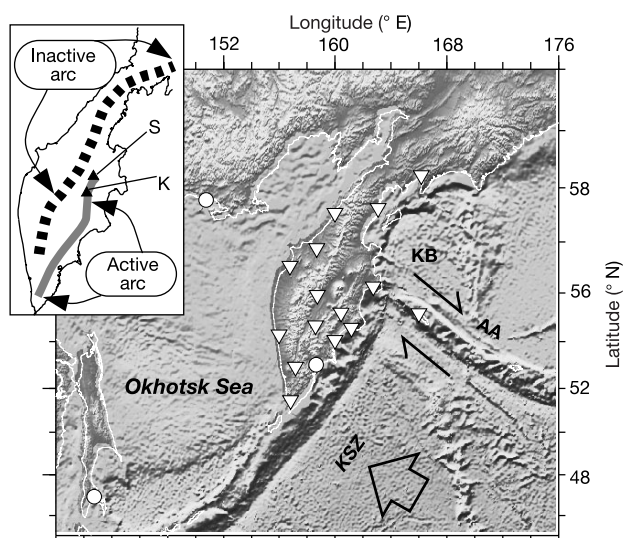


Figure 1 An overview map of the Aleutian–Kamchatka junction (AKJ), showing the tectonic setting and the observational system used. Shaded topography outlines the mid-Kamchatka (Sredinny) range of Kamchatka that is a locus of past island-arc volcanism. Open symbols denote permanent (circle) and temporary (triangle) seismic observatories. The large open arrow shows the Pacific plate motion direction, filled half-arrows denote a region of distributed strike-slip deformation. Inset, outlines of two volcanic chains of Kamchatka, and locations of the Klyuchevskoy volcanic group (K) and the Sheveluch volcano (S). Geographic and tectonic features marked are the Aleutian Arc (AA), the Kamchatka subduction zone (KSZ) and the Komandorsky basin (KB).

configuration of recently and presently subducted oceanic lithosphere, and the mantle flow regime at the AKJ. The fate of the relict slab beneath northern Kamchatka is unknown, as no deep earthquakes occur north of the AKJ¹². A relict slab beneath the western Aleutian islands¹⁵ could accommodate the change in plate-boundary type at the AKJ by folding over the corner, but petrologic data¹⁶ suggest a gap in subducted lithosphere between the AKJ and Adak island in the central Aleutians.

Tomographic imaging of the upper-mantle compressional velocity (V_p) structure in the AKJ region with body waves^{17,18} is limited by a scarcity of stations and seismicity north of the AKJ, and poor crossing-ray coverage from teleseismic events in the uppermost mantle. Surface-wave dispersion tomography is better suited for imaging shear velocity (V_s) in the key depth range (50–150 km) for lithospheric detachment. The primary data for our tomographic V_s

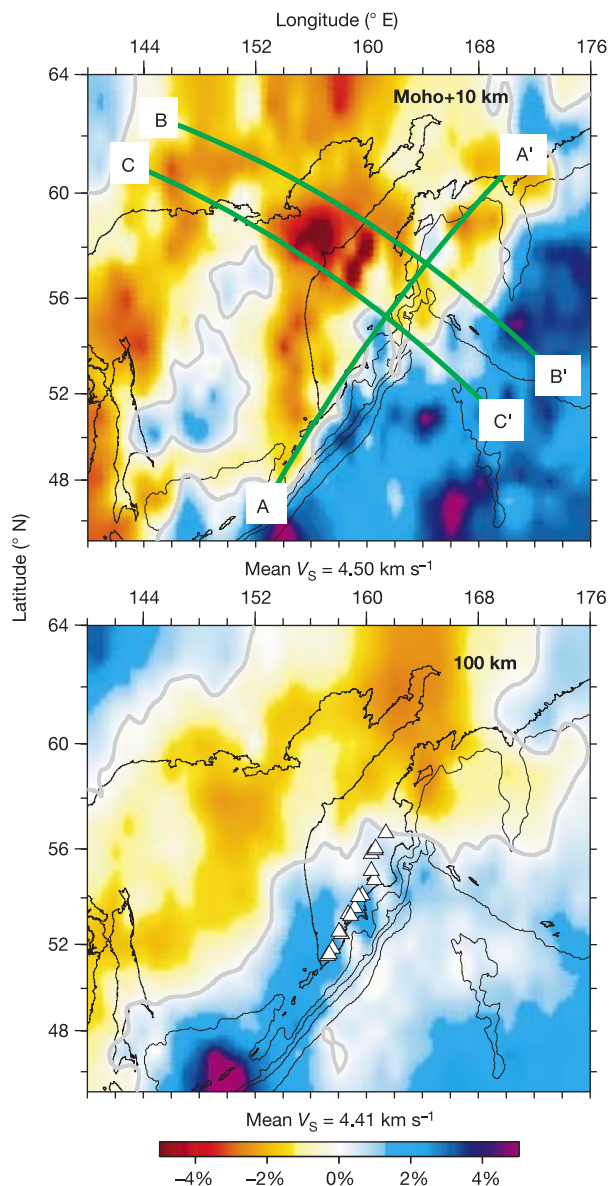


Figure 2 Maps of seismic velocity distribution. Deviations from the mean regional value are shown for the depth of 100 km (bottom panel) and for 10 km beneath the crust–mantle transition (top panel). Variations in crustal thickness imply an uneven surface in the latter graph. Light grey lines show contours of mean regional velocity. Coastlines and outlines of submarine bathymetry are shown for reference. Top, green lines show tracks of the vertical profiles presented in Fig. 3. Bottom, volcanoes with post-1900 eruptions are plotted as open triangles.

model consists of a large set of dispersion curves for broadband Rayleigh and Love wave group and phase velocities^{19,20,21}. Because long propagation paths in global tomographic studies can limit lateral resolution, we added complementary dispersion data (group velocity measurements for 1,432 Rayleigh and 649 Love source–receiver paths) from a year-long deployment of portable seismic stations in Kamchatka³⁰. Period-specific dispersion maps were found with “diffraction tomography”²², a technique based on a physical model of the surface wave Fresnel zone. These dispersion maps were subjected to nonlinear Monte-Carlo inversion²³ for an ensemble of acceptable V_S models of the crust and the uppermost mantle on a $1^\circ \times 1^\circ$ grid across the region of study.

Descending cold mantle lithosphere should be characterized by a positive V_S anomaly with thin (~ 200 km) tabular geometry. The joint inversion of short- and intermediate-period group speeds with long-period phase speed data greatly reduces the ensemble variance of the acceptable V_S models at individual gridpoints, and affords tighter lateral resolution in the shallow mantle. Vertical resolution is improved by constraining crustal structure with body-wave reverberations²⁴. The multiplicity of sources and receivers means that lateral resolution for our study region is better than the global

average. Vertical resolution experiments establish that the high- V_S signature of a subducting slab can be estimated reliably to depths of at least 200 km.

Figures 2, 3 and 4a show that a fast V_S anomaly delineates the descending Pacific plate beneath southern Kamchatka (profile

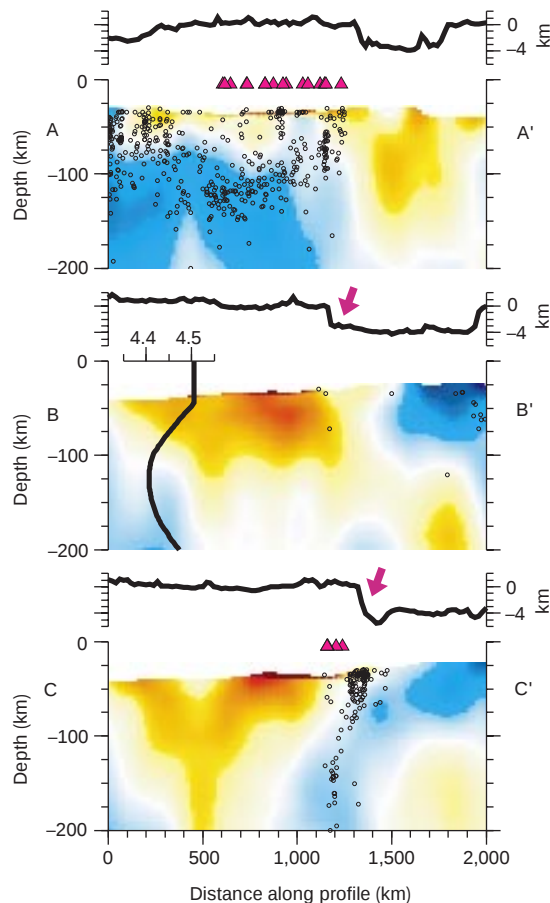


Figure 3 Vertical cross-sections through the shear velocity model with the same colour scale as in Fig. 2, vertical exaggeration 1:5. Colours indicate anomalies in S-wave velocity relative to the regional one-dimensional distribution shown in profile B–B'. The shallow part of the model is removed to accentuate the details of the deeper structures. Surface topography/bathymetry along profile tracks is plotted (in km) above each profile. Earthquake hypocentres within 20 km of the profile plane are shown by small circles. Volcanoes with post-1900 eruptions within 100 km of the profile are shown by triangles. Trench location is marked by the arrow of profiles B–B' and C–C'. Subducted slab of the Komandorsky basin lithosphere would have been to the left of the arrow on profile B–B'. In the panel showing profile B–B', the inset horizontal axis shows shear wave velocity, V_S , in km s^{-1} .

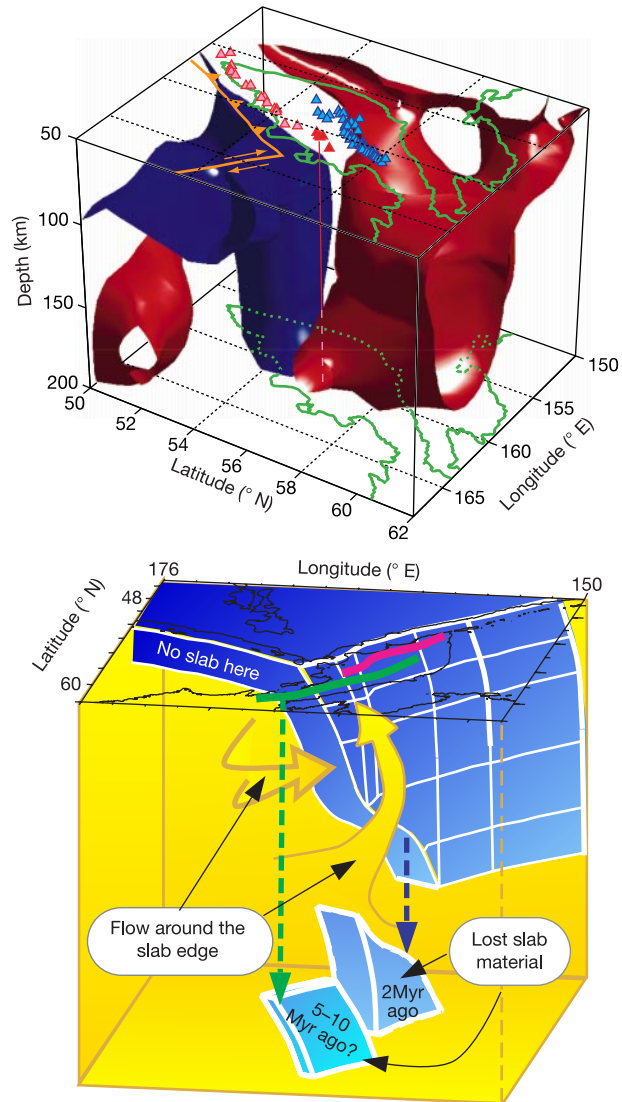


Figure 4 Structure and geodynamics of the AKJ. **a**, Isosurface representation of the V_S model beneath the AKJ, in which the model was laterally smoothed with a gaussian filter ($\sigma = 100$ km) to highlight the dominant large-scale features. The blue surface ($+0.8\%$) represents the high seismic velocity oceanic lithosphere subducting beneath southern Kamchatka. The red surface (-2.2%) reflects low seismic velocity material beneath northern Kamchatka. The coastline (green) and the modern plate boundary (salmon-coloured arrows and line) are identified. Pink and red triangles are locations of active volcanoes where red triangles correspond to the Klyuchevskoy and Sheveluch volcanic groups located above the edge of the subducted slab; blue triangles are the locations of Holocene volcanoes in the presently inactive mid-Kamchatka volcanic belt; the red line (solid and dashed) shows the position of the volcanic group marked by red triangles relative to the structure at depth. In the late Miocene and into the Pliocene, volcanic activity in this belt extended up to 60° N (ref. 13). **b**, A depiction of the subducting Pacific plate (blue). Outlines of the active (magenta) and inactive (green) volcanic arcs are shown. Fragments of the younger Komandorsky lithosphere (lighter blue) and the older Pacific lithosphere (deeper blue) has detached from the lithospheric mantle beneath the AKJ (as shown by dashed arrows), at 5–10 and 2 Myr ago, respectively. This allowed warmer asthenospheric material (yellow arrows) to flow both around the subducting slab and upwards towards the surface.

C–C'), weakens northward, and terminates beneath the AKJ (profile A–A'). The anomaly is particularly weak where seismicity gets shallower beneath the AKJ. A slow anomaly beneath the Okhotsk Sea is contiguous with a deeper slow anomaly under the Komandorsky basin (profile B–B').

These V_S images, historical seismicity¹², and earlier V_P tomography^{17,18} provide evidence that the current oceanic lithosphere of the Pacific plate terminates at the AKJ. Neither the V_S nor the previous V_P tomographic images detect relict slab material beneath northern Kamchatka from the approximately 10 Myr ago subduction of the Komandorsky basin⁴. The results from V_S tomography, in combination with other geophysical and geological evidence, show that the upper mantle beneath northern Kamchatka has undergone major changes since 10 Myr ago and that subduction dynamics beneath Kamchatka is far from steady-state. We argue that this instability in the upper mantle involves at least two episodes of slab loss (Fig. 4b). Specifically, we propose that in the first episode, which occurred between 5 Myr and 10 Myr ago, a relict slab created by westward subduction of the Komandorsky basin lithosphere detached. In the second episode, since 2 Myr ago, a fragment of the present Pacific plate detached and sank into the mantle.

We associate the first episode of slab loss, that is, detachment of Komandorsky lithosphere, with the latest (Pliocene-epoch to the Quaternary period) stage of the volcanism in northern Kamchatka. The absence of an aseismic slab combined with a progressive change from island-arc volcanism to rift-like volcanism at about 5 Myr ago^{13,14} suggest thermal rejuvenation of the mantle lithosphere beneath northern Kamchatka since that time. Alkalic volcanism consistent with low partial-fraction pressure–release melting is a likely consequence of asthenospheric upwelling^{3,13,14}. Petrologic evidence for a “slab melt” component in volcanic rocks from the now-extinct northern Kamchatka arc²⁵ may be associated with the remelting of dacite-veined peridotite within the supra-slab mantle wedge, a lithology reported in northern Kamchatka mantle xenoliths²⁶. Low values of V_S imaged in the asthenosphere beneath northern Kamchatka further support the idea of the influx of warm asthenospheric material towards the surface.

The second episode of slab detachment, involving the loss of an edge fragment of the Pacific plate within the past 2 Myr ago or so, is inferred on the basis of Pliocene–Pleistocene tectonic motions, as well as the volcanism near the AKJ⁵. The presently subaerial central Kamchatka depression near the AKJ lay beneath sea level in the Pliocene epoch³. Active volcanism in the Klyuchevskoy volcanic cluster lies above the slab edge itself (Fig. 4a), and dates back only about 50 kyr ago²⁷. The current output of Klyuchevskoy volcano exceeds that of any other island-arc volcano⁶. Its lavas are characterized by high temperatures of equilibration²⁸, consistent with pressure–release melting induced by a transient lofting of the slab edge^{5,6}. Slab-edge lofting causes the inland shift of the island-arc volcanism at the AKJ relative to the volcanic arc further south¹². A pulse of subduction-derived volatiles is an alternative cause of intensified volcanism, but contradicts geochemical evidence for modest sediment input²³ and long mantle-wedge residence times of volatiles²⁹. In addition, the northern part of the fast V_S anomaly that we associate with the Pacific lithosphere is very weak (~1%), and its border is highly irregular. This weakening probably reflects the loss of material from the side edge of the subducting slab, through dynamic or gravitational instability⁷ or through erosion by mantle flow around the corner of the subduction zone³. Such flow would have been obstructed if the Komandorsky lithosphere had not detached. Further evidence for such flow, together with some material ascent towards the surface beneath northwestern Kamchatka, comes from seismic anisotropy^{5,30}, and from an exceptionally slow V_S anomaly to the northwest of the AKJ that is apparently connected to a deeper body of slow material beneath the Komandorsky basin (Fig. 4a). This flow regime may be responsible for a ‘slab melt’, or adakite, petrologic signature at Sheveluch, the north-

ernmost active volcano in Kamchatka, through partial melting of the slab edge¹⁶. It is likely that the erosion of the slab edge is an continuing process.

Our findings emphasize the importance of slab detachment and loss as an agent in the evolution of convergent margins. They also illustrate the complexity of dynamical, petrological and tectonic processes that occur near slab edges. □

Received 19 March; accepted 8 July 2002; doi:10.1038/nature00973.

1. Zonenshain, L. P., Kuzmin, M. I. & Natapov, L. M. *Geology of the USSR: A Plate-Tectonic Synthesis* (AGU, Geodynamics Series No. 21, American Geophysical Union, Washington DC, 1990).
2. Geist, E. L. & Scholl, D. W. Application of continuum models to deformation of the Aleutian island arc. *J. Geophys. Res.* **97**, 4953–4967 (1992).
3. Erlich, E. N. & Gorshkov, G. S. (eds) *Quaternary volcanism and tectonics in Kamchatka*. *Bull. Volcanol.* **42**, 1–298 (1979).
4. Honthaas, C., Bellon, H., Kepezhinskias, P. K. & Maury, R. C. New ⁴⁰K–⁴⁰Ar dates for the Cretaceous–Quaternary magmatism of Northern Kamchatka (Russia). *C. R. Acad. Sci. Paris Serie II* **320**, 197–204 (1995).
5. Park, J. et al. in *Plate Boundary Zones* (eds Stein, S. & Freymuller, J.) 295–324 (AGU, Geodynamics Series No. 30, American Geophysical Union, Washington DC, 2002).
6. Kersting, A. B. & Arculus, R. J. Klyuchevskoy volcano, Kamchatka, Russia: The role of high-flux recharged, tapped, and fractionated magma chamber(s) in the genesis of high-Al₂O₃ from high-MgO basalt. *J. Petrol.* **35**, 1–41 (1994).
7. Davies, J. H. & von Blanckenburg, F. Slab breakoff: A model of lithosphere detachment and its test in the magmatism and deformation of collisional orogens. *Earth Planet. Sci. Lett.* **129**, 85–102 (1995).
8. Wortel, M. J. R. & Spakman, W. Subduction and slab detachment in the Mediterranean–Carpathian region. *Science* **290**, 1910–1917 (2000).
9. Lallemand, S. E., Font, Y., Bijwaard, H. & Kao, H. New insights on 3-D plates interaction near Taiwan from tomography and tectonic implications. *Tectonophysics* **335**, 229–253 (2001).
10. Kosarev, G. L. et al. Seismic evidence for a detached Indian lithospheric mantle beneath Tibet. *Science* **283**, 1306–1309 (1999).
11. Seber, D., Barazangi, M., Ibenbrahim, A. & Demnati, A. Geophysical evidence for lithospheric delamination beneath the Alboran Sea and Rif-Betic mountains. *Nature* **379**, 785–790 (1996).
12. Gorbato, A., Kostaglodov, V., Suarez, G. & Gordeev, E. Seismicity and structure of the Kamchatka subduction zone. *J. Geophys. Res.* **102**, 17883–17898 (1997).
13. Fedorov, P. I. & Shapiro, M. N. Neogene volcanics of the Kamchatka isthmus and geodynamics of the Aleutian–Kamchatka junction. *Geotectonics* **32**, 122–137 (1998).
14. Volynets, O. N. Geochemical types, petrology and genesis of late Cenozoic volcanic rocks from the Kurile–Kamchatka Island–Arc system. *Int. Geol. Rev.* **36**, 373–405 (1994).
15. Boyd, T. M. & Creager, K. C. The geometry of Aleutian subduction: Three-dimensional seismic imaging. *J. Geophys. Res.* **96**, 2267–2291 (1991).
16. Yagodinski, G. M. et al. Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges. *Nature* **409**, 500–504 (2001).
17. Van der Hilst, R. D., Engdahl, E. R., Spakman, W. & Nolet, G. Tomographic imaging of subducted lithosphere below northwest Pacific island arcs. *Nature* **353**, 37–43 (1991).
18. Gorbato, A., Widiyantoro, A., Fukao, Y. & Gordeev, E. Signature of remnant slabs in the North Pacific from P-wave tomography. *Geophys. J. Int.* **142**, 27–36 (2000).
19. Ritzwoller, M. H. & Levshin, A. L. Eurasian surface wave tomography: group velocities. *J. Geophys. Res.* **103**, 4839–4878 (1998).
20. Ekstrom, G., Tromp, J. & Larson, E. W. F. Measurements and global models of surface waves propagation. *J. Geophys. Res.* **102**, 8137–8157 (1997).
21. Trampert, J. & Woodhouse, J. H. Global phase velocity maps of Love and Rayleigh waves between 40 and 150 s period. *Geophys. J. Int.* **122**, 675–690 (1995).
22. Ritzwoller, M. H., Shapiro, N. M., Barmin, M. P. & Levshin, A. L. Global surface wave diffraction tomography. *J. Geophys. Res.* (in the press).
23. Shapiro, N. M. & Ritzwoller, M. H. Monte-Carlo inversion of broad-band surface wave dispersion for a global shear velocity model of the crust and upper mantle. *Geophys. J. Int.* (in the press).
24. Levin, V. et al. Crust and upper mantle of Kamchatka from teleseismic receiver functions. *Tectonophysics* (in the press).
25. Kepezhinskias, P. K. et al. Trace element and Sr–Nd–Pb isotopic constraints on a three-component model of Kamchatka Arc petrogenesis. *Geochim. Cosmochim. Acta* **61**, 577–600 (1997).
26. Kepezhinskias, P. K., Defant, M. J. & Drummond, M. S. Progressive enrichment of island arc mantle by melt–peridotite interaction inferred from Kamchatka xenoliths. *Geochim. Cosmochim. Acta* **60**, 1217–1229 (1996).
27. Braitseva, O. A., Melekstev, I. V., Ponomareva, V. V. & Sulerzhitsky, L. D. Ages of calderas, large explosive craters and active volcanoes in the Kuril–Kamchatka region, Russia. *Bull. Volcanol.* **57**, 383–402 (1995).
28. Ozerov, A. Y. The evolution of high-alumina basalts of the Klyuchevskoy volcano, Kamchatka, Russia, based on microprobe analysis of mineral inclusions. *J. Volcanol. Geotherm. Res.* **95**, 65–79 (2000).
29. Turner, S., McDermott, E., Hawkesworth, C. & Kepezhinskias, P. A U-series study of lavas from Kamchatka and the Aleutians: Constraints on source composition and melting processes. *Contrib. Mineral. Petrol.* **133**, 217–234 (1998).
30. Peyton, V. et al. Mantle flow at a slab edge: Seismic anisotropy in the Kamchatka region. *Geophys. Res. Lett.* **28**, 379–382 (2001).

Acknowledgements

We acknowledge help with seismological fieldwork from the Kamchatka Experimental-Methodical Seismological Department of Russian Academy of Sciences and assistance

from IRIS Data Management Center in data and archive management. This work was funded by the National Science Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to V.L. (e-mail: vadim.levin@yale.edu).

A primitive fish close to the common ancestor of tetrapods and lungfish

Min Zhu* & Xiaobo Yu†

* Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, PO Box 643, Beijing 100044, China

† Department of Biological Sciences, Kean University, Union, New Jersey 07083, USA

The relationship of the three living groups of sarcopterygians or lobe-finned fish (tetrapods, lungfish and coelacanth) has been a matter of debate^{1–5}. Although opinions still differ, most recent phylogenies suggest that tetrapods are more closely related to lungfish than to coelacanth^{6–10}. However, no previously known fossil taxon exhibits a concrete character combination approximating the condition expected in the last common ancestor of tetrapods and lungfish—and it is still poorly understood how early sarcopterygians diverged into the tetrapod lineage (Tetrapodomorpha)⁷ and the lungfish lineage (Dipnomorpha)⁷. Here we describe a fossil sarcopterygian fish, *Styloichthys changae* gen. et sp. nov., that possesses an eyestalk and which exhibits the character combination expected in a stem group close to the last common ancestor of tetrapods and lungfish. *Styloichthys* from the Lower Devonian of China bridges the morphological gap between stem-group sarcopterygians (*Psarolepis* and *Achoania*)¹⁰ and basal tetrapodomorphs/basal dipnomorphs. It provides information that will help in the study of the relationship of early sarcopterygians, and which will also help to resolve the tetrapod–lungfish divergence into a documented sequence of character acquisition.

Sarcopterygii (Romer, 1955)

Styloichthys gen. nov.

Diagnosis. A sarcopterygian characterized by relatively large pores (often spoon-shaped and arranged in parallel grooves) on the cosmine surface, a jagged margin between ethmosphenoid and otoccipital shields, an otoccipital with a wide flat ventral surface carrying no vestibular fontanelle, and a lower jaw with a ventrally protruding flange formed by prearticular and meckelian bone. *Styloichthys* differs from *Psarolepis* and *Achoania* in having a lyre-shaped trajectory of the supraorbital canal, a fenestra ventralis, small internasal cavities and three coronoids in the lower jaw. *Styloichthys* differs from tetrapodomorphs and dipnomorphs in having an eyestalk, a slender postorbital pila, and a cosmine surface with relatively large pores.

Type species. *Styloichthys changae* sp. nov.

Etymology. Generic name referring to the presence of postorbital pila (Greek *stilo*, pillar; *ichthys*, fish). Specific name in honour of M.-M. Chang for her contributions to palaeoichthyology.

Holotype. V8142.1, an anterior cranial portion, Institute of Vertebrate Paleontology and Paleoanthropology, Beijing.

Age and locality. Early Devonian (late Lochkovian), Xitun Formation, Qujing, East Yunnan, China.

Remarks. The new form (Figs 1 and 2) is represented by 20 anterior cranial portions (V8142.1–20), 7 posterior cranial portions (V8142.21–27), 18 lower jaws (V8143.1–18), one maxillary (V8143.19), 4 cheek bone plates (V8143.20–23), 5 cleithra (V8143.24–28) and 4 clavicles (V8143.29–32), all showing a unique cosmine surface with relatively large pores (though smaller than those in *Psarolepis* and *Achoania*^{10–12}). The assignment of these specimens to the same form is further supported by matching morphology between the anterior and posterior cranial portions (Fig. 1a–f) and between the upper-jaw and lower-jaw structures (Figs 1c, i, and 2b, d).

Description. The most remarkable feature of *Styloichthys* is a recessed teardrop-shaped eyestalk area immediately behind the optic canal (Fig. 1e, g, h). This unfinished area has a well-defined natural margin as indicated by the surrounding periosteal lining which dips slightly into the unfinished recess, and corresponds to the eyestalk attachment area recently reported in basal bony fishes (*Psarolepis*, *Achoania* and *Ligulalepis* formerly known as AMF101607)^{10,13,14}. Ventral to this eyestalk area, a cup-shaped depression serves for eye muscle attachment. *Styloichthys* also resembles *Psarolepis* and *Achoania* in having a postorbital pila (though splinter-like and much more slender) that forms a bridge between the top of the basiptyergoid process and the side of the braincase wall.

Styloichthys is quite close to *Youngolepis*¹⁵ (a basal dipnomorph) in having an anteriorly positioned ethmoidal articulation, a fenestra ventralis and small internasal cavities, while other endocranial features seem intermediate between *Youngolepis* and *Psarolepis/Achoania*, such as a relatively broad suborbital ledge, a large notochordal pit, and a well-developed intracranial joint with a processus connectens giving a slightly convex profile to the posterior face of the ethmosphenoid (Fig. 1c, e, g, h). The well-ossified otoccipital (Fig. 1d, f) carries a large oval concave area for the basicranial muscle, a large keyhole-shaped bipartite facet for the hyomandibular and a large facet for the first infrapharyngeal branchial, but no vestibular fontanelle (also absent in some coelacanth and megalichthyid tetrapodomorphs such as *Cladrosymblema*¹⁶).

Styloichthys resembles *Youngolepis* and, to a lesser extent, *Powichthys*¹⁷ (a basal dipnomorph) and *Kenichthys*¹⁸ (a basal tetrapodomorph), in dermal bone features such as an inward and downward bending snout, a small independent premaxillary with high antero-medial portion and low postero-lateral portion (indicated by the outline of the premaxillary area and the sutural position of ethmoidal commissural canal), a large and long parasphenoid, a pineal foramen at the anterior margin of the parietals, a lyre-shaped trajectory of the supraorbital canal, groups of pores piercing the smooth surface between the relatively large cosmine pores, and a postparietal laterally flanked by more than two bones and posteriorly overlapped by the extrascapular. The compound cheek bone plate (corresponding to squamosal + quadratojugal + preopercular), maxillary, clavicle, and cleithrum with a tripartite scapulo-coracoid (Fig. 2) are also similar to those in *Youngolepis*¹⁹.

The lower jaw (Figs 1i, 2c, d) has three coronoids as indicated by three semilunar pockets along a shallow groove between the dentary and the prearticular, but is unique in having a very large adductor fossa (making up more than 50% of the lower jaw length), a convex ventral flange formed by prearticular and meckelian bone and protruding beyond the ventral border of infradentaries, and a prearticular carrying a shagreen of minute denticles dorsally and undulating parallel ridges ventrally.

To explore the phylogenetic position of *Styloichthys*, we modified the data matrix in ref. 10 by adding the *Styloichthys* codings and revising 21 codings for *Ligulalepis* and *Onychodus* based on newly available information^{13,20}. Phylogenetic analysis²¹ (see Supplementary Information) using the modified data matrix of 27 taxa and 158 characters yields 24 trees showing the same major groupings as ref. 10. *Styloichthys* is placed in a trichotomy with Dipnomorpha and

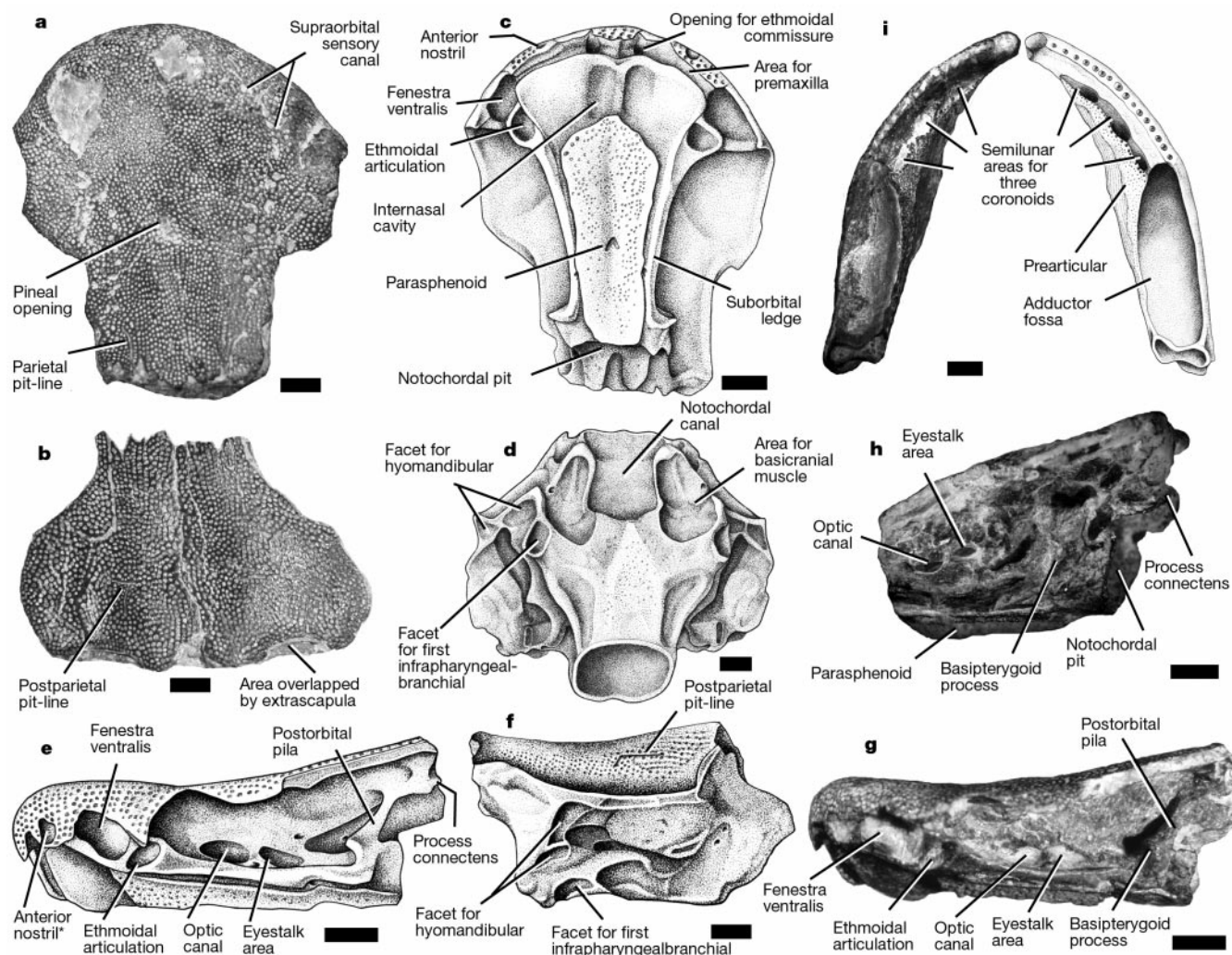


Figure 1 *Styloichthys changae* gen. et sp. nov., a primitive fossil fish close to the last common ancestor of tetrapods and lungfish. Anterior cranial portion (**a**, V8142.5; **c**, **e**, **g**, V8142.1, holotype; **h**, V8142.4), posterior cranial portion (**b**, V8142.24; **d**, **f**, V8142.21) and lower jaw (**i**, V8143.1). **a**, **b**, Dorsal view (photographs), showing the lyre-shaped trajectory of the supraorbital sensory canal, the jagged margin between the ethmosphenoid and otoccipital shields, and the relatively large pores on the cosmine surface. **c**, **d**, Ventral view (drawings), showing the large parasphenoid, large notochordal

pit and canal, and the wide otoccipital with no vestibular fontanelle. **e**–**h**, Lateral view (drawings and photographs), showing the eyestalk area, the postorbital pila, the processus connectens and the facets for the hyomandibular and the first infrapharyngealbranchial. Left anterior nostril (marked with asterisk in **e**) is reconstructed from well-preserved right anterior nostril, partially visible in **c**. **i**, Dorsal view (photograph with mirror-image drawing) showing the semilunar areas for three coronoids and the large adductor fossa. Scale bar, 2 mm.

Tetrapodomorpha in the 50% majority rule consensus tree, but shows up more often at the base of Tetrapodomorpha + Dipnomorpha (9 trees) than at other positions (6 trees at the base of Tetrapodomorpha, 3 trees at the base of Dipnomorpha and 6 trees within Dipnomorpha). This result, together with the distribution of the eyestalk, the postorbital pila and the large-pore cosmine, suggests that *Styloichthys* is best regarded as the stem group of Tetrapodomorpha + Dipnomorpha directly below the node representing their last common ancestor (Fig. 3).

The phylogenetic position of *Styloichthys* and its unique character combination have wide implications for studying the relationship of early sarcopterygians. First, the presence of postorbital pila and large-pore cosmine in *Styloichthys* provides evidence that these are primitive features for sarcopterygians rather than autapomorphies uniting *Psarolepis* and *Achoania*, which form successive plesions, and not a monophyletic clade, in the stem-group segment of the Sarcopterygii. Together with the eyestalk, these features must have been independently lost at least twice within the Sarcopterygii (once in the coelacanth + onychodonts clade, and once in the last

common ancestor of tetrapods and lungfish).

Second, *Styloichthys* allows more precise statements to be made about the sequence of character acquisition leading to the tetrapod–lungfish divergence. For instance, the straight course of the supraorbital canal in *Psarolepis*, *Achoania* and coelacanth (for example, *Miguashaia* and *Whiteia*)^{22,23}, the 5 coronoids in *Psarolepis* and *Achoania* (undescribed material), and the variable pattern of 4–5 coronoids in coelacanth²² all resemble the actinopterygian condition²⁴ and differ from the dipnomorph + tetrapodomorph condition. *Styloichthys* documents the first appearance of the lyre-shaped trajectory of the supraorbital canal and the three coronoid condition found in dipnomorphs + tetrapodomorphs, thereby placing the origin of these features in the internode between the coelacanth + onychodonts clade and *Styloichthys*.

Third, *Styloichthys* provides a crucially positioned new outgroup to examine character polarity and phylogenetic relationship among dipnomorphs and tetrapodomorphs. For instance, in 6 of the 9 trees placing *Styloichthys* at the base of Tetrapodomorpha + Dipnomorpha, *Powichthys* is aligned with porolepiforms rather

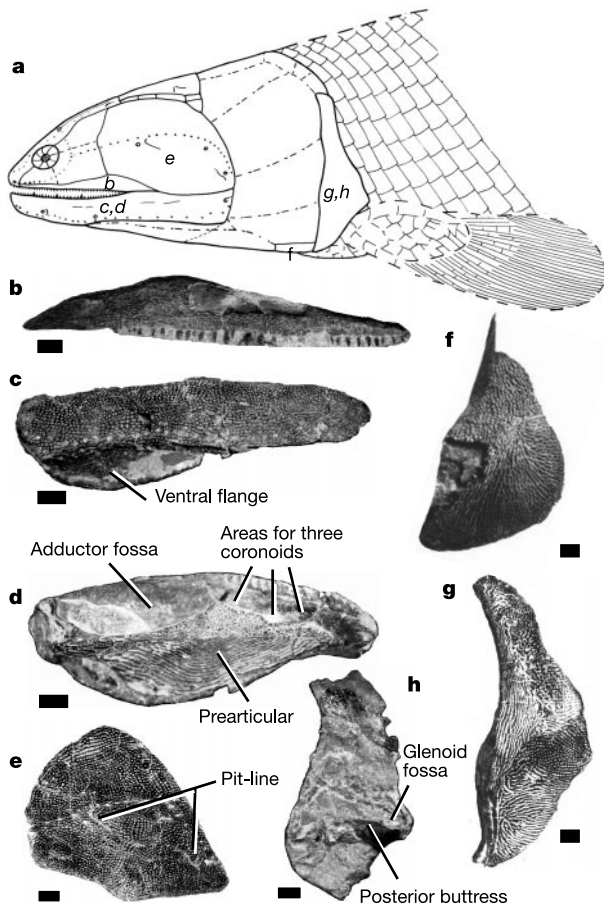


Figure 2 Skull and shoulder girdle of *Styloichthys*. **a**, Reconstructed lateral view showing positions of dermal bones. **b**, Right maxillary (V8143.19) in lateral view. **c**, Right lower jaw (V8143.3) in lateral view, showing the ventrally protruding flange. **d**, Left lower jaw (V8143.1) in medial view showing the semilunar areas for three coronoids and the large adductor fossa. **e**, Left compound cheek plate (V8143.20) corresponding to squamosal + quadratojugal + preopercular. **f**, Left clavicle (V8143.29). **g**, Left cleithrum (V8143.24). **h**, Right cleithrum (V8143.25) in medial view, showing structures of the scapulocoracoid. Scale bar, 2 mm.

than with dipnoforms (in the sense of ref. 8). This new tentative position of *Powichthys* (Fig. 3) may offer a more parsimonious explanation for the distribution of the intracranial joint, absent in *Youngolepis*¹⁵, *Diabolepis*²⁵ and dipnoans²⁶ (lungfish), but present in both porolepiforms (including *Powichthys*) and tetrapodomorphs (with subsequent parallel loss in the fish–tetrapod transition)²⁷. Similarly, in the context of *Styloichthys*, porolepiform-like features found in rhizodonts (such as a short snout, widely separated eyes, and extratemporalis)^{28,29} as well as similarities among *Youngolepis*, *Powichthys* and *Diabolepis*, porolepids and early osteolepids³⁰, can be more confidently explained as primitive features of dipnomorphs + tetrapodomorphs.

Last, *Styloichthys* sheds light on the possible tempo and mode of evolutionary changes surrounding the tetrapod–lungfish divergence. The Lochkovian age of *Styloichthys* and *Youngolepis* suggests that tetrapodomorphs and dipnomorphs diverged very rapidly from their common ancestor. Although the dermal bones and lower jaws in *Styloichthys* approach the tetrapodomorph–dipnomorph condition, its endocranium exhibits a mixture of features found in *Psarolepis/Achoania* and in basal tetrapodomorphs/basal dipnomorphs. This suggests that changes in dermal bones (including the lower jaw) outpaced changes in endocranial features, and probably set the stage for divergent morphological specializations

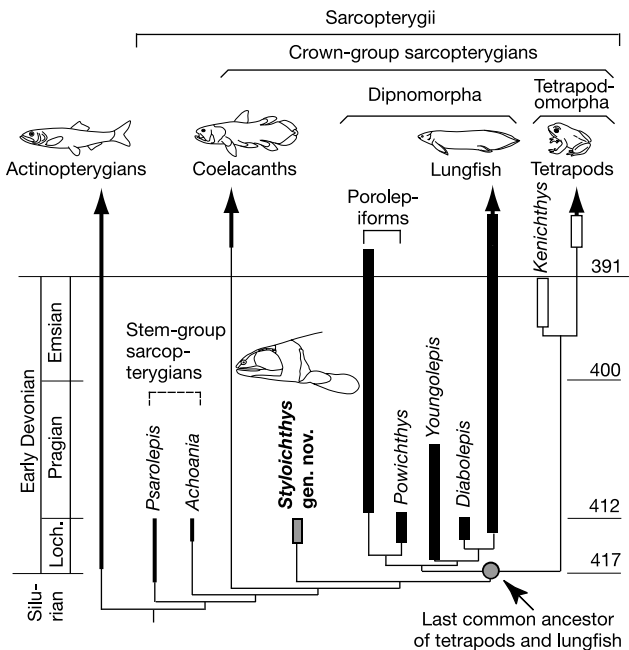


Figure 3 Stratogram showing *Styloichthys* as the stem group of the clade Dipnomorpha + Tetrapodomorpha. (The position of *Styloichthys* is based on 9 out of 24 trees, the position of *Powichthys* is based on 6 trees; tree length, 328; consistency index, 0.549; homoplasy index, 0.451; retention index, 0.759; rescaled consistency index, 0.416; see Supplementary Information for details). Numbers along the right margin represent millions of years before present.

leading respectively to the tetrapod lineage and the lungfish lineage. □

Received 27 December 2001; accepted 9 May 2002; doi:10.1038/nature00871.

- Rosen, D. E., Forey, P. L., Gardiner, B. G. & Patterson, C. Lungfishes, tetrapods, paleontology and plesiomorphy. *Bull. Am. Mus. Nat. Hist.* **167**, 159–276 (1981).
- Schultze, H.-P. Comparison of hypotheses on the relationships of sarcopterygians. *Syst. Biol.* **43**, 155–173 (1994).
- Janvier, P. *Early Vertebrates* (Oxford Univ. Press, Oxford, 1996).
- Zardoya, R. et al. Searching for the closest living relative(s) of tetrapods through evolutionary analyses of mitochondrial and nuclear data. *Mol. Biol. Evol.* **15**, 506–517 (1998).
- Zhu, M. & Schultze, H.-P. in *Major Events in Early Vertebrate Evolution* (ed. Ahlberg, P. E.) 289–314 (Taylor & Francis, London, 2001).
- Maisey, J. G. Head and tails: a chordate phylogeny. *Cladistics* **2**, 201–256 (1986).
- Ahlberg, P. E. A re-examination of sarcopterygian interrelationships, with special reference to the Porolepiformes. *Zool. J. Linn. Soc.* **103**, 241–287 (1991).
- Cloutier, R. & Ahlberg, P. E. in *Interrelationships of Fishes* (eds Stiassny, M. L., Parenti, L. R. & Johnson, G. D.) 445–480 (Academic, San Diego, 1996).
- Meyer, A. Molecular evidence on the origin of tetrapods and the relationships of the coelacanth. *Trends Ecol. Evol.* **35**, 93–101 (1995).
- Zhu, M., Yu, X. & Ahlberg, P. A primitive sarcopterygian fish with an eyestalk. *Nature* **410**, 81–84 (2001).
- Yu, X. A new porolepiform-like fish, *Psarolepis romeri*, gen. et. sp. nov. (Sarcopterygii, Osteichthyes) from the Lower Devonian of Yunnan, China. *J. Vert. Paleontol.* **18**, 261–274 (1998).
- Zhu, M., Yu, X. & Janvier, P. A primitive fossil fish sheds light on the origin of bony fishes. *Nature* **397**, 607–610 (1999).
- Basden, A. M. & Young, G. C. A primitive actinopterygian neurocranium from the early Devonian of southeastern Australia. *J. Vert. Paleontol.* **21**, 754–766 (2001).
- Basden, A. M. et al. The most primitive osteichthyan braincase? *Nature* **403**, 185–188 (2000).
- Chang, M. M. *The Braincase of Youngolepis, a Lower Devonian Crossosporian from Yunnan, South-western China* Thesis, Stockholm Univ. (1982).
- Fox, R. C. et al. A new osteolepiform fish from the Lower Carboniferous Raymond Formation, Drummond Basin, Queensland. *Mem. Queensl. Mus.* **38**, 97–221 (1995).
- Jessen, H. L. Lower Devonian Porolepiformes from the Canadian Arctic with special reference to *Powichthys thorsenianus* Jessen. *Palaeontogr. A* **167**, 180–214 (1980).
- Chang, M. M. & Zhu, M. A new Middle Devonian osteolepidid from Qujing, Yunnan. *Mem. Assoc. Australas. Paleontol.* **15**, 183–198 (1993).
- Chang, M. M. in *Early Vertebrates and Related Problems of Evolutionary Biology* (eds Chang, M. M., Liu, Y. H. & Zhang, G. R.) 355–378 (Science, Beijing, 1991).
- Long, J. On the relationships of *Psarolepis* and the onychodontiform fishes. *J. Vert. Paleontol.* **21**, 815–820 (2001).
- Swofford, D. L. *PAUP: Phylogenetic Analysis Using Parsimony ver. 3.1.1* (Illinois Natural History Survey, Champaign, Illinois, 1993).

22. Forey, P. L. *History of the Coelacanth Fishes* (Chapman & Hall, London, 1998).
23. Cloutier, R. in *Devonian Fishes and Plants of Miguasha, Quebec, Canada* (eds Schultze, H.-P. & Cloutier, R.) 227–247 (Pfeil, Munich, 1996).
24. Gardiner, B. G. The relationships of the palaeoniscid fishes, a review based on new specimens of *Mimia* and *Moythomasia* from the Upper Devonian of Western Australia. *Bull. Br. Mus. Nat. Hist. Geol.* **37**, 173–428 (1984).
25. Chang, M. M. *Diabolepis* and its bearing on the relationships between porolepiforms and dipnoans. *Bull. Mus. Natl. Hist. Nat. Ser.* **4** **17**, 235–268 (1995).
26. Schultze, H.-P. in *The Biology and Evolution of Lungfishes* (eds Bemis, W. E., Burggren, W. W. & Kemp, N. E.) *J. Morph.* **1** (Suppl.), 39–74, (1987).
27. Ahlberg, P. E. & Johanson, Z. Osteolepiforms and the ancestry of tetrapods. *Nature* **395**, 792–794 (1998).
28. Long, J. A new rhizodontiform fish from the Early Carboniferous of Victoria, Australia, with remarks on the phylogenetic position of the group. *J. Vert. Paleontol.* **9**, 1–17 (1989).
29. Johanson, Z. & Ahlberg, P. E. A complete primitive rhizodont from Australia. *Nature* **394**, 569–572 (1998).
30. Chang, M. M. & Yu, X. Re-examination of the relationship of Middle Devonian osteolepids—fossil characters and their interpretations. *Am. Mus. Novit.* **3189**, 1–20 (1997).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

We thank M. M. Chang and P. E. Ahlberg for advice and discussions, K. S. Thomson for comments and corrections, M. Yang for artwork, J. Zhang for photographic work, and X. Lu for specimen preparation. This work was supported by the Special Funds for Major State Basic Research Projects of China, the Chinese Foundation of Natural Sciences, and the US National Geographic Society. X.Y. thanks Kean University for faculty research and development support.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.Z. (e-mail: zhumin@ht.rol.cn.net).

Increasing dominance of large lianas in Amazonian forests

Oliver L. Phillips*, **Rodolfo Vásquez Martínez†**, **Luzmila Arroyo‡§**, **Timothy R. Baker***, **Timothy Killeen‡§||**, **Simon L. Lewis*¶**, **Yadvinder Malhi¶**, **Abel Monteagudo Mendoza†‡#**, **David Neill☆☆**, **Percy Núñez Vargas#**, **Miguel Alexiades††**, **Carlos Cerón‡‡**, **Anthony Di Fiore§§**, **Terry Erwin|||**, **Anthony Jardim§**, **Walter Palacios☆☆**, **Mario Saldias§** & **Barbara Vinceti¶**

* Centre for Biodiversity and Conservation, School of Geography, University of Leeds LS2 9JT, UK

† Jardín Botánico de Missouri, Jaén, Peru

‡ Missouri Botanical Garden, St Louis, Missouri 63166-0299, USA

§ Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia

|| Conservation International Washington DC 20036, USA

¶ The School of Earth Environmental and Geographical Sciences, University of Edinburgh EH9 3JU, UK

Herbario Vargas, Universidad San Antonio Abad del Cusco, Cusco, Peru

☆☆ Fundación Jatun Sacha; ** Missouri Botanical Garden; ‡‡ Herbario QAP, Escuela de Biología de la Universidad Central del Ecuador, Quito, Ecuador

†† New York Botanical Garden, Bronx, New York 10458, USA

§§ Department of Anthropology, New York University, New York 10003, USA

||| Natural History Museum, Smithsonian Institution, Washington DC 20560, USA

Ecological orthodoxy suggests that old-growth forests should be close to dynamic equilibrium, but this view has been challenged by recent findings that neotropical forests are accumulating carbon^{1,2} and biomass^{3,4}, possibly in response to the increasing atmospheric concentrations of carbon dioxide^{5,6}. However, it is unclear whether the recent increase in tree biomass has been

accompanied by a shift in community composition. Such changes could reduce or enhance the carbon storage potential of old-growth forests in the long term. Here we show that non-fragmented Amazon forests are experiencing a concerted increase in the density, basal area and mean size of woody climbing plants (lianas). Over the last two decades of the twentieth century the dominance of large lianas relative to trees has increased by 1.7–4.6% a year. Lianas enhance tree mortality and suppress tree growth⁷, so their rapid increase implies that the tropical terrestrial carbon sink may shut down sooner than current models suggest^{8–10}. Predictions of future tropical carbon fluxes will need to account for the changing composition and dynamics of supposedly undisturbed forests.

Recent field studies^{1,2,3} indicate that old-growth tropical forests are absorbing 1–2 Gt C yr⁻¹, but the mechanisms and stability of the tropical carbon sink, and its implications for the ecology of tropical vegetation, are highly uncertain. Shifts in functional composition and biodiversity are expected as a result of climate changes and increased CO₂ (refs 11, 12) but so far there is no evidence of widespread compositional change in old-growth forests. This absence of evidence might imply evidence of absence—or it could simply reflect our failure to monitor adequately forest behaviour, or even to examine existing data across sufficient spatial and temporal scales. Lianas in particular are ignored in forest inventories and models alike, in spite of their key functional roles. As structural parasites, lianas exert a much greater ecological effect than their size suggests, representing less than 5% of tropical forest biomass but up to 40% of leaf productivity¹³. They also suppress tree growth and encourage tree mortality, and affect the competitive balance among trees by disproportionately infesting some taxa and suppressing the regeneration and growth of non-pioneers⁷. Climbers respond strongly to increased CO₂ concentrations^{14,15} and benefit from disturbance^{7,16,17}, and a biome-wide trend to increased tree turnover rates has been detected in old-growth forests¹⁸ so increases in liana densities might be anticipated¹⁹. Here we assemble several unique,

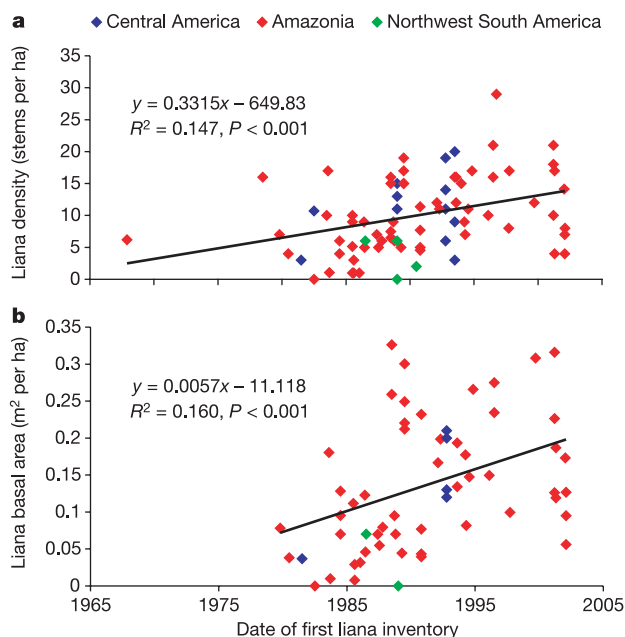


Figure 1 Structural importance of lianas over 10 cm in diameter in each neotropical site as a function of date of first inventory. **a**, Liana stem density in stems ha⁻¹; **b**, liana basal area in m² ha⁻¹. 'Central America' is Panama and tropical countries to the north; 'Northwest South America' is the Chocó bioregion, west of the Andes; 'Amazonia' is the Amazon river basin and contiguous forested zones of Guyana and eastern Brazil. Linear regressions are fitted to the Amazonian data.

Table 1 **Modelled changes in liana density and basal area**

Dependent variable	Source of variation				
	A. Environment only		B. Environment plus time		
	Adjusted R^2 (%)	F-value, d.f.	Δ Adjusted R^2 (%)	F-value, d.f.	T-value and significance for T in model
Liana stem density (ha^{-1})	35.3	3.83*, 26	+13.2	5.89***, 26	+2.93**
Liana basal area ($\text{m}^2 \text{ha}^{-1}$)	27.2	2.94*, 26	+21.1	7.06***, 26	+4.42***
Number of lianas/number of trees	34.1	4.36*, 26	+12.4	6.66***, 26	+3.93***
Basal area of lianas/basal area of trees	34.6	3.60*, 26	+21.7	6.59***, 26	+4.80***
Mean liana basal area (cm^2)	0.3	1.04, 26	+5.5	1.54, 26	+1.55

*, $0.05 > P \geq 0.01$; **, $0.01 > P \geq 0.001$; ***, $P < 0.001$.

Multiple regression of large liana density and basal area (BA) at each initial census, modelled as a function of the environment, and the environment plus time. Environmental variables measured at North Peru, South Peru and Bolivian sites include climate (mean annual rainfall, seasonality), soil chemistry (pH, Ca, K, Mg, Na, P, Al), soil particle size distribution (sand, silt, clay), and hydrology (drainage and risk of water-logging). Soil variables were converted to principal components before multiple regression; time variable (T, in years) is the decimal date in which the liana parameter was first recorded in the plot. See Methods for details.

long-term, multi-regional data sets of liana and tree populations. We use them to test both the general hypothesis that the composition of old-growth tropical forests is changing over large scales, and the specific prediction that lianas are benefiting. We analyse 47 interior-forest sites in four Amazonian regions (North Peru, South Peru, Bolivia and Ecuador) where we are monitoring all woody plants of over 10 cm in diameter in 1-ha plots, and include published data from a further 37 neotropical sites. We find that the density and the basal area of large lianas have increased substantially over the last two decades of the twentieth century. The same trends are observed however liana populations are analysed.

First, we plot liana dominance as a function of each site's first inventory data (Fig. 1). There is broad scatter, reflecting forest variation and large sampling error at the plot scale¹⁷, but a

significant trend for late-censused neotropical sites to have greater liana dominance than early-censused sites. Analysis of covariance (ANCOVA) shows that this is not an artefact of spatial changes in sampling intensity—in models with lianas as the dependent variable, and region and year as independent variables, the year contributes significantly to both neotropical and Amazon liana stem density ($F_{1,69} = 15.30$, $P < 0.001$; $F_{1,53} = 14.37$, $P < 0.01$) and to neotropical and Amazon liana basal area ($F_{1,49} = 8.96$, $P < 0.01$; $F_{1,43} = 6.99$, $P < 0.02$). The increases in lianas as a function of first inventory dates are also unlikely to be artefacts of a change through time in the environmental characteristics of the forests sampled, because the date of first inventory contributes significantly to statistical models of liana density and basal area even after accounting for edaphic and climatic effects (Table 1).

Second, we analysed a different data set: the changes that

Table 2 **Linear trends in the structural importance of large lianas**

a Amazonian sites with a monitoring period of over 5 yr

Parameter	Annual rate of change in parameter (mean $\pm$ 95% CI)	Annual rate of change (proportion of site initial value)	Annual rate of change (proportion of site final value)
Liana stem density (ha^{-1})	+0.22 $\pm$ 0.11 ($n = 28$)	+4.03 $\pm$ 2.56%	+1.78 $\pm$ 0.82%
Liana basal area ($\text{m}^2 \text{ha}^{-1}$)	+3.72 $\pm$ 1.16 $\times 10^{-3}$ ($n = 28$)	+4.58 $\pm$ 2.60%	+2.40 $\pm$ 0.62%
Liana stems as a fraction of tree stems	+3.45 $\pm$ 2.10 $\times 10^{-4}$ ($n = 28$)	+3.27 $\pm$ 2.10%	+1.70 $\pm$ 0.87%
Liana basal area as a fraction of tree basal area	+1.19 $\pm$ 0.53 $\times 10^{-4}$ ($n = 28$)	+4.05 $\pm$ 2.30%	+2.07 $\pm$ 0.71%
Mean basal area per liana stem (cm^2)	+1.03 $\pm$ 0.97 ($n = 27$)	+0.96 $\pm$ 0.66%	+0.65 $\pm$ 0.65%

b Amazonian sites with a monitoring period of over 5 yr analysed by region

Parameter	ANOVA, test of hypothesis of a regional cluster effect	Annual rate of change in parameter (mean $\pm$ 95% CI) for each region
Liana stem density (ha^{-1})	$F = 0.84$, $P = 0.48$	N. Peru +0.34 $\pm$ 0.25 ($n = 7$) S. Peru +0.19 $\pm$ 0.22 ($n = 12$) Bolivia +0.06 $\pm$ 0.41 ($n = 5$) Ecuador +0.28 $\pm$ 0.32 ($n = 4$)
Liana basal area ($\text{m}^2 \text{ha}^{-1}$)	$F = 0.64$, $P = 0.60$	N. Peru +0.0049 $\pm$ 0.0022 ($n = 7$) S. Peru +0.0029 $\pm$ 0.0025 ($n = 12$) Bolivia +0.0042 $\pm$ 0.0029 ($n = 5$) Ecuador +0.0033 $\pm$ 0.0033 ($n = 4$)
Liana stems as a fraction of tree stems	$F = 0.95$, $P = 0.43$	N. Peru +6.34 $\pm$ 4.75 $\times 10^{-4}$ ($n = 7$) S. Peru +2.20 $\pm$ 3.91 $\times 10^{-4}$ ($n = 12$) Bolivia +2.10 $\pm$ 7.89 $\times 10^{-4}$ ($n = 5$) Ecuador +4.30 $\pm$ 5.70 $\times 10^{-4}$ ($n = 4$)
Liana basal area, as a fraction of tree basal area	$F = 0.62$, $P = 0.61$	N. Peru +1.79 $\pm$ 0.82 $\times 10^{-4}$ ($n = 7$) S. Peru +0.97 $\pm$ 1.21 $\times 10^{-4}$ ($n = 11$) Bolivia +0.80 $\pm$ 1.81 $\times 10^{-4}$ ($n = 5$) Ecuador +1.11 $\pm$ 1.46 $\times 10^{-4}$ ($n = 4$)
Mean basal area per liana stem (cm^2)	$F = 0.15$, $P = 0.93$	N. Peru +0.66 $\pm$ 2.92 ($n = 7$) S. Peru +0.96 $\pm$ 1.83 ($n = 11$) Bolivia +1.61 $\pm$ 1.90 ($n = 5$) Ecuador +1.20 $\pm$ 2.49 ($n = 4$)

ANOVA, analysis of variance. CI, confidence interval.

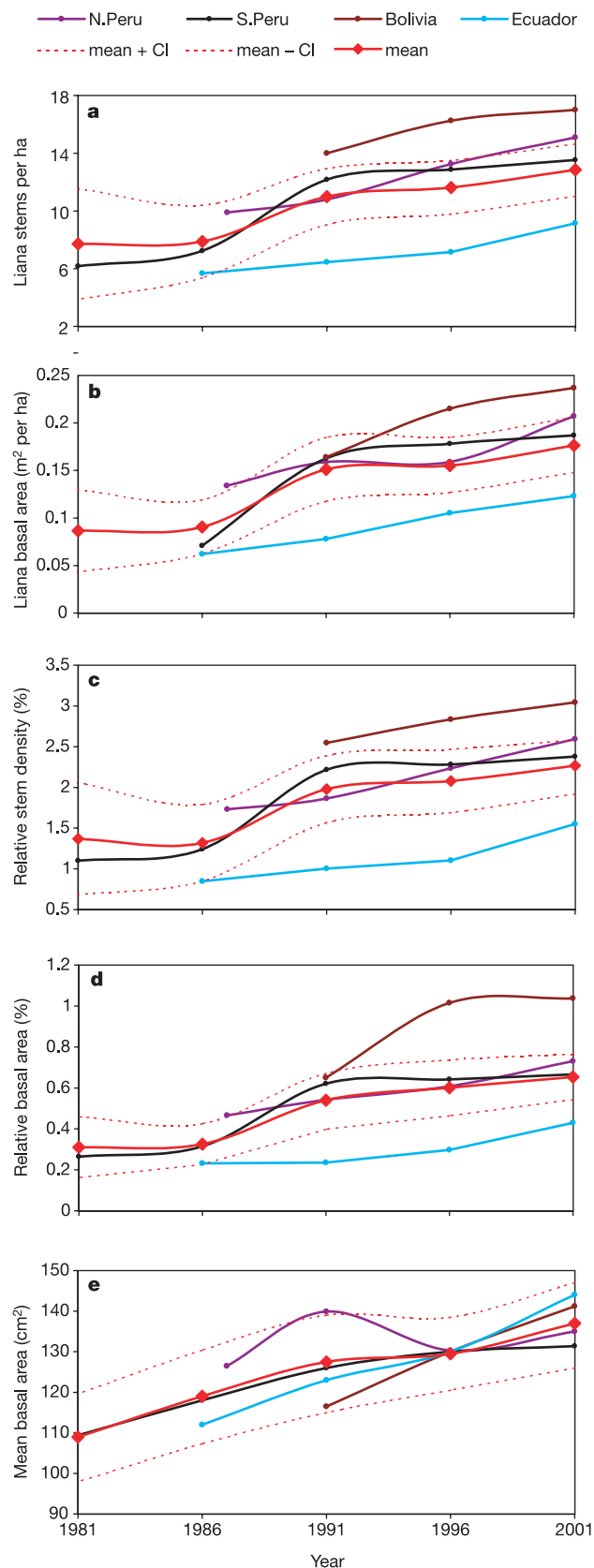


Figure 2 Changes through time of the importance of lianas over 10 cm in diameter in western Amazonia. **a**, Liana stem density in stems ha^{-1} ; **b**, liana basal area in $\text{m}^2 \text{ha}^{-1}$; **c**, relative liana density as a percentage of tree stems; **d**, relative liana basal area as a percentage of tree basal area; **e**, mean basal area of each liana stem in cm^2 . Graphs show 5-yr running means with 95% confidence intervals, with values plotted separately for North Peru, South Peru, Bolivia and Ecuador.

occurred through time within multi-census sites. Here we find consistently strong positive changes in measures of liana dominance (1.7–4.6% per year) and stem size (0.6–1.0% per year), evident across all regions (Table 2).

Finally, we assembled all available data from our four west Amazonian regions, that is single-census and multi-census data, to create running means across sites, and find similarly consistent patterns over the last two decades (Fig. 2). The year-on-year increase in mean values of large lianas is not driven by a few atypical sites or by the intrinsic liana richness of any one region, but is rather a general phenomenon across all five ecological parameters and all four regions (ANCOVA, with year as the continuous variable and region as fixed factor, shows that year contributes ($P < 0.01$) for all 20 combinations except for mean stem size in North Peru ($P > 0.05$)).

We examined the underlying dynamics of lianas in all plots censused three or more times, and find that the rates of both large liana growth and large liana loss have increased (comparing annual rate of basal area gain for lianas interval 1 versus interval 2, $t = -3.10$, $n = 21$, $P < 0.01$; annual rate of basal area loss for lianas interval 1 versus interval 2, $t = -2.66$, $n = 21$, $P < 0.02$), whereas growth rates have consistently exceeded loss rates (annual rate of basal area gain versus loss for lianas ≥ 10 cm in diameter for interval 1: $t = 4.19$, $n = 21$, $P < 0.001$; and for interval 2: $t = 2.38$, $n = 21$, $P < 0.05$). Thus the net increase in liana basal area was driven by high liana growth rates that have increased through time, and occurred in spite of an acceleration in the rate of liana mortality.

Tree basal area increased by $0.34 \pm 0.20\%$ a year in the 1980s and 1990s in Amazonia³ but the increase in liana values has been much more rapid: the relative importance of large lianas has approximately doubled over a similar period across all sites (Fig. 2) and the annual rates of increase in liana density and BA within sites exceeded the rate of increase in tree BA by an order of magnitude (Table 2a).

Mature Amazonian forest plots have therefore undergone substantial change in functional composition. The increase in lianas has been concerted in the sense that it has occurred simultaneously over a wide spatial, climatic and edaphic range. We have shown that this is not an effect of changes in the kinds of forests sampled through time, and other potential artefactual explanations can also be ruled out (Supplementary Information). We can explore the generality of the changes indicated by the permanent plot data set with an independent data set from 70 old-growth lowland forests across the Neotropics. These are single-census 0.1-ha forest samples surveyed between 1971 and 1997, in which every tree and liana ≥ 2.5 cm in diameter was inventoried once (see Methods), thus sampling plants with basal area as low as about 6% of the smallest plants in our permanent plots. After controlling for climate and soil effects, the neotropical 0.1-ha data set confirms the temporal trends seen in the permanent plot data set: liana popu-

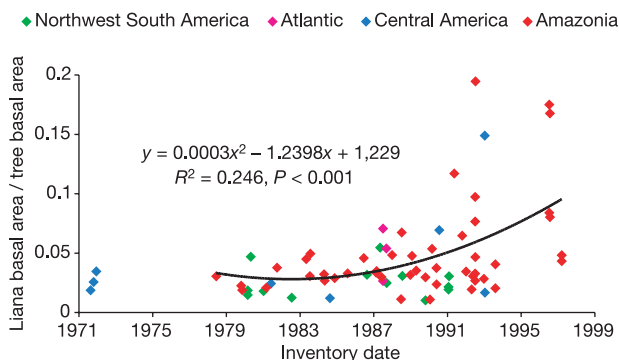


Figure 3 Relative dominance of lianas in neotropical 0.1-ha plots as a function of inventory date. Liana basal area is shown as a fraction of tree basal area for all stems over 2.5 cm in diameter. A polynomial curve is fitted for the Amazonian sites.

lations have become more dense, liana basal area has increased, the relative dominance of lianas has increased, and the size of individual lianas has increased (Fig. 3, and Supplementary Information Table 3). The degree of internal consistency within and between data sets across differing sample unit sizes, target variables, minimum plant sizes, climatic regimes, edaphic conditions, regional locations and spatial scales is a critical factor in assessing confidence in the changes. The results demonstrate a substantial increase in the density and relative dominance of lianas in western Amazonia, and available evidence suggests that this change in the structural and functional composition of forests has been even more widespread.

We asked what is driving this change? If regional climates have changed in western Amazonia or across the neotropics that could provide an explanation, but we failed to find convincing evidence for this (Supplementary Information). The documented increase in CO₂ concentrations is another possibility, because lianas respond strongly to CO₂ fertilization over the historical range of concentrations¹⁵. The direct impacts of CO₂ on photosynthesis may drive nonlinear compositional responses—for example, the relative rate of stimulation on liana growth may be particularly strong in deep shade¹⁵, enhancing the likelihood of lianas reaching the sunlit canopy. Effects of increasing CO₂ on climber growth^{14,15} could also be magnified in a positive feedback loop with the simultaneous increase in tree turnover¹⁸ (more rapid turnover favours gap-specialist lianas which in turn accelerate tree mortality).

The increased liana load in trees may have a major impact on the Amazon carbon sink. The biomass of lianas themselves is usually small^{7,20}, but lianas can substantially suppress tree biomass^{7,17}. We examined the relationship between liana infestation and subsequent tree mortality in the 13 sites where we have accurate records of liana–host relationships. Here, liana infestation of tree biomass is associated with a $39.6 \pm 31.3\%$ excess risk of tree mortality. The expected annual rate of increase in the amount of tree biomass mortality is estimated as $1.64 \pm 1.38\%$, the product of the liana-associated excess tree mortality risk and the annual rate of increase in large lianas per unit of tree biomass in the sites. However climate models suggest that in east and south Amazonia moisture supply will become more seasonal^{9,10}, conditions which may favour lianas (Supplementary Information), so synergisms between climate change and increasing liana densities could magnify the impact of either process alone. Better understanding of these risks will require intensive field research to improve the liana-on-tree mortality functions and to begin including lianas within full tropical forest vegetation models and coupled carbon cycle/climate models.

The increase in liana density and biomass is the first evidence for a widespread functional shift in old-growth tropical forests. Regardless of the impacts on the carbon cycle, this has important implications for the biodiversity of tropical forests. First, if the increase is driven by increasing CO₂ concentrations it implies that the extensive tropical forests of the Cretaceous and Tertiary periods when CO₂ concentrations peaked at $>2,000$ p.p.m. (ref. 21) may have differed radically from today's in structure and function. Second, lianas and trees are differentiated phylogenetically²² and by distinctive pollination and dispersal ecologies²³, so changed relative densities has knock-on consequences for conservation of plants and animals. Third, increased liana density has the potential to alter tree species composition because climber impacts on trees vary with host phylogeny and ecology^{7,16}. Finally, the change in composition has direct societal and economic impacts, because lianas are valued less than trees by forest communities²⁴ and are major silvicultural pests for the tropical timber industry⁷. □

Methods

Study sites of 1 ha

We censused large liana populations in 47 sites spanning the climatic and edaphic gradients of western Amazonia²⁵ (Supplementary Information Appendix). 1-ha permanent plots were sited since as early as 1979 in old-growth forest and recensused every

2–5 yr, most recently in 2002 (eight sites) and 2001 (19 sites), yielding up to 19 yr of growth and dynamics data. Plot locations were constrained by the need for reasonable access (<10 km to nearest road or navigable river) and long-term protection, but are otherwise sited randomly or haphazardly within landform strata and are unbiased by sylvigenetic state^{3,4}. The straight-line distance between all pairs of plot centroids ranges by four orders of magnitude (0.2–2,380 km), but intersite distance had no effect on liana change metrics (Supplementary Information). Distances to edges are 0.1 to ~ 3.6 km, where 'edge' is defined as a previously forested location where there has been an anthropogenic impact creating a canopy gap of ≥ 0.5 ha the effects of which are still apparent at the time of the first census. Edges were formed by farmers, research stations, tourist lodges and logging activities. We also searched the literature for published single-census large liana inventories in neotropical forests with $\geq 1,500$ mm rain, including all except montane/cloud forests, small fragments ($<1,000$ ha), or with known human disturbance to forest structure.

Liana and tree measurements

Tree measurements and analysis follow RAINFOR protocols²⁵ (<http://www.geog.leeds.ac.uk/projects/rainfor/>). For lianas the diameter of each independently rooted climbing stem rooted within each plot and potentially over 10 cm wide was measured at a height of 1.3 m ($d_{1.3}$: all sites except 14, 15 and 25) and at the widest point within 2.5 m of the ground ($d_{\max}$: all sites except 23, 24, 34 and 35), or at both points for most censuses (40 sites). We used the $d_{1.3}:d_{\max}$ ratio of individual lianas to determine $d_{\max}$ of lianas at censuses where only $d_{1.3}$ was recorded. This procedure gives unbiased estimates, because the ratio of $d_{1.3}$ to $d_{\max}$ is independent of $d_{\max}$ (mean $\pm$ 95% confidence interval, CI, of r , the correlation between $d_{1.3}/d_{\max}$ and $d_{\max} = 0.007 \pm 0.080$, $n = 28$). For both measurement methods we calculated the number of lianas and total basal area (the sum of cross-sectional stem areas) at each census. Lianas support greater biomass and productivity than trees of equal diameter^{7,13,26}, so stem density and basal area are not equivalent measures of functional importance between life-forms, but for each life-form basal area provides a close approximation to biomass^{3,26,27}. For reasons of brevity we only report results based on $d_{1.3}$ measurements. Results based on $d_{\max}$ are presented in the Supplementary Information Tables 1 and 2 and are essentially equivalent.

Climate and soil data

Climate data were sourced directly from local records, or indirectly from interpolated maps^{25,28}, to derive mean annual rainfall and seasonality (consecutive months averaging <100 mm rain) for every neotropical site. At Bolivia and South and North Peru sites, soils were sampled from up to ten randomly chosen locations within each plot at 0–15 or 0–20 cm depth; samples were bulked, dried, and subsampled. Analysis followed standard ISRIC procedures²⁹. Most sites have been visited at least five times, allowing assessment of hydrologic conditions on a scale of one (permanently water-logged) to ten (excessively draining).

Change in permanent plots

For within-site and within-region change analyses (Table 2) we included all interior-forest old-growth sites with at least two full censuses of trees and lianas over 10 cm in diameter more than 5 yr apart. For between-site change analyses (Fig. 1, Table 1) we included interior-forest old-growth plots with at least one full census of trees and lianas over 10 cm in diameter. To minimize effects of any asymmetric sampling of environments through time, we first excluded permanent swamp and white sand sites and then used principal components ordination analysis (PCA) to describe the major gradients in normalized and standardized soil variables in Peruvian and Bolivian sites, and then applied multiple regression to test the effects of the PCA factors, climate variables and the time variable on forest structure (Table 1), repeating the procedure with neotropical 0.1-ha sites (Supplementary Information Table 3). We report multiple-regression models with greatest adjusted- r^2 values and control for climate and edaphic effects by computing the variance explained by inventory date after accounting for the variance explained by environmental variables.

For the displays of change within- and between-sites (Fig. 2), we included interior-forest old-growth plots with at least one full census of trees and lianas. We used linear interpolation between each census to estimate structural values within sites and then derived cross-site 5-yr running means to smooth the effects of site-switching.

For comparisons of liana growth and loss within plots we split each monitoring period into intervals of similar length (first interval, 5.9 ± 0.7 yr; second interval, 6.0 ± 0.9 yr), so that turnover rates could be compared directly while controlling for possible effects of interval length on estimated mortality and growth rates.

Study sites of 0.1 ha

Inventories were completed by A. Gentry (58 sites) (<http://www.mobot.org/MOBOT/Research/gentry/transects.html>), R.V.M. and O.L.P. (6), and T.K., L.A. and M.S. (6). All scandent lianas and hemiepiphytes with $d_{\max}$ over 2.5 cm and non-climbing stems with $d_{1.3}$ over 2.5 cm were measured. See ref. 30 for detailed descriptions. All available neotropical old-growth forest samples with more than 1,500 mm rain were included, except montane/cloud forests, small fragments (less than 1,000 ha), or sites with known human disturbance to forest structure (extra-Amazonian sites also exclude island hurricane-impacted forests).

Received 25 March; accepted 30 May 2002; doi:10.1038/nature00926.

- Grace, J. *et al.* Carbon dioxide uptake by an undisturbed tropical rain-forest in Southwest Amazonia, 1992–1993. *Science* **270**, 778–780 (1995).
- Malhi, Y. *et al.* Carbon dioxide transfer over a Central Amazonian rain forest. *J. Geophys. Res. Atmos.* **103**, 31593–31612 (1998).

3. Phillips, O. L. *et al.* Changes in the carbon balance of tropical forest: evidence from long-term plots. *Science* **282**, 439–442 (1998).
4. Phillips, O. L. *et al.* Changes in the biomass of tropical forests: evaluating potential biases. *Ecol. Appl.* **12**, 576–587 (2002).
5. Prentice, I. C. *et al.* in *Intergovernmental Panel on Climate Change Third Assessment Report, Climate Change 2001: The Scientific Basis* Ch. 3 (Cambridge Univ. Press, Cambridge, UK, 2001).
6. Malhi, Y. & Grace, J. Tropical forests and atmospheric carbon dioxide. *Trends Ecol. Evol.* **15**, 332–337 (2000).
7. Schnitzer, S. A. & Bongers, F. The ecology of lianas and their role in forests. *Trends Ecol. Evol.* **17**, 223–230 (2002).
8. Chambers, J. Q., Higuchi, N. & Tribuzy, E. S. Carbon sink for a century. *Nature* **410**, 429–429 (2001).
9. Cox, P. M. *et al.* Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
10. White, A., Cannell, M. G. R. & Friend, A. D. CO₂ stabilisation, climate change and the terrestrial carbon sink. *Glob. Change Biol.* **6**, 817–833 (2000).
11. Condit, R., Hubbell, S. P. & Foster, R. B. Assessing the response of plant functional types to climatic change in tropical forests. *J. Veg. Sci.* **7**, 405–416 (1996).
12. Körner, C. Biosphere responses to CO₂ enrichment. *Ecol. Appl.* **10**, 1590–1619 (2000).
13. Hegarty, E. E. & Caballé, G. in *The Biology of Vines* (eds Putz, F. E. & Mooney, H. A.) 313–336 (Cambridge Univ. Press, Cambridge, UK, 1991).
14. Condon, M. A., Sasek, T. W. & Strain, B. R. Allocation patterns in two tropical vines in response to increased atmospheric CO₂. *Funct. Ecol.* **6**, 680–685 (1992).
15. Granados, J. & Körner, C. In deep shade, elevated CO₂ increases the vigour of tropical climbing plants. *Glob. Change Biol.* (in the press).
16. Pérez-Salici, D. R., Sork, V. L. & Putz, F. E. Lianas and trees in Amazonian Bolivia. *Biotropica* **33**, 34–37 (2001).
17. Laurance, W. F. *et al.* Rain forest fragmentation and the structure of Amazonian liana communities. *Ecology* **82**, 105–116 (2001).
18. Phillips, O. L. & Gentry, A. H. Increasing turnover through time in tropical forests. *Science* **263**, 954–958 (1994).
19. Phillips, O. L. The changing ecology of tropical forests. *Biodivers. Cons.* **6**, 291–311 (1997).
20. Putz, F. E. Liana biomass and leaf-area of a terra firme forest in the Rio Negro basin, Venezuela. *Biotropica* **15**, 185–189 (1983).
21. Retallack, G. J. A 300 million year record of atmospheric carbon dioxide from fossil plant cuticles. *Nature* **411**, 287–290 (2001).
22. Gentry, A. H. in *The Biology of Vines* (eds Putz, F. E. & Mooney, H. A.) 3–49 (Cambridge Univ. Press, Cambridge, UK, 1991).
23. Gentry, A. H. in *The Biology of Vines* (eds Putz, F. E. & Mooney, H. A.) 393–423 (Cambridge Univ. Press, Cambridge, UK, 1991).
24. Phillips, O. L. & Gentry, A. H. The useful plants of Tambopata, Peru. II: Additional hypothesis testing in quantitative ethnobotany. *Econ. Bot.* **47**, 33–43 (1993).
25. Malhi, Y. *et al.* An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). *J. Veg. Sci.* (in the press).
26. Gerwing, J. J. & Lopes Farias, D. Integrating liana abundance and forest stature into an estimate of total aboveground biomass for an eastern Amazonian forest. *J. Trop. Ecol.* **16**, 327–335 (2000).
27. Brown, S. *Estimating Biomass and Biomass Change of Tropical Forests: a Primer* (Food and Agriculture Organisation Forestry Paper 134, Rome, 1997).
28. Sombroek, W. G. Spatial and temporal patterns of Amazon rainfall: consequences for the planning of agricultural occupation and the protection of primary forests. *Ambio* **30**, 388–396 (2001).
29. van Reeuwijk, L. P. (ed.) *Procedures for Soil Analysis*, Tech. Pap. 9, 5th edn (International Soil Reference and Information Centre, FAO, Rome, 1995).
30. Phillips, O. L. & Miller, J. *Global Patterns of Plant Diversity: Alwyn H. Gentry's Forest Transect Data Set* (Missouri Botanical Garden, St Louis, in the press).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

We acknowledge the contributions of more than 50 field assistants in Peru, Ecuador and Bolivia, the residents of Constanica, Infierno, La Torre, Mishana and Florida, as well as logistical support from Instituto Nacional de Recursos Naturales (INRENA), Amazon Center for Environmental Education and Research (ACEER), Cuzco Amazónico Lodge, Explorama Tours SA, Instituto de Investigaciones de la Amazonía Peruana (IIAP), Parque Nacional Noel Kempff, Peruvian Safaris SA, Universidad Nacional de la Amazonía Peruana, and Universidad Nacional de San Antonio Abad del Cusco. Field research was supported by the EU Fifth Framework Programme (RAINFOR), the UK Natural Environment Research Council, the National Geographic Society, the American Philosophical Society, the National Science Foundation, the WWF-U.S./ Garden Club of America, Conservation International, the MacArthur and Mellon Foundations, US-AID, the Max-Planck Institute for Biogeochemistry and the Royal Society (Y.M.). The manuscript benefited from comments by C. Körner and N. Pitman. We are indebted to the late A.H. Gentry for helping to make this work possible.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to O.P. (e-mail: o.phillips@geog.leeds.ac.uk).

Involvement of DARPP-32 phosphorylation in the stimulant action of caffeine

Maria Lindskog*, Per Svenningsson†, Laura Pozzi*‡, Yong Kim†, Allen A. Fienberg†, James A. Bibb†‡, Bertil B. Fredholm§, Angus C. Nairn†, Paul Greengard† & Gilberto Fisone*

* Department of Neuroscience, Karolinska Institutet, 17177 Stockholm, Sweden
† Laboratory of Molecular and Cellular Neuroscience, The Rockefeller University, New York, New York 10021, USA

§ Department of Physiology and Pharmacology, Karolinska Institutet, 17177 Stockholm, Sweden

‡ Present addresses: Department of Neuroscience, "Mario Negri" Institute for Pharmacological Research, Milan, Italy (L.P.); Department of Psychiatry, UT Southwestern Medical Center, Dallas, Texas, USA (J.A.B)

Caffeine has been imbibed since ancient times in tea and coffee, and more recently in colas. Caffeine owes its psychostimulant action to a blockade of adenosine A_{2A} receptors¹, but little is known about its intracellular mechanism of action. Here we show that the stimulatory effect of caffeine on motor activity in mice was greatly reduced following genetic deletion of DARPP-32 (dopamine- and cyclic AMP-regulated phosphoprotein of relative molecular mass 32,000)². Results virtually identical to those seen with caffeine were obtained with the selective A_{2A} antagonist SCH 58261. The depressant effect of the A_{2A} receptor agonist, CGS 21680, on motor activity was also greatly attenuated in DARPP-32 knockout mice. In support of a role for DARPP-32 in the action of caffeine, we found that, in striata of intact mice, caffeine increased the state of phosphorylation of DARPP-32 at Thr 75. Caffeine increased Thr 75 phosphorylation through inhibition of PP-2A-catalysed dephosphorylation, rather than through stimulation of cyclin-dependent kinase 5 (Cdk5)-catalysed phosphorylation, of this residue. Together, these studies demonstrate the involvement of DARPP-32 and its phosphorylation/dephosphorylation in the stimulant action of caffeine.

Striatal medium spiny neurons have an important role in the control of voluntary movements. A large subpopulation of these neurons project to the substantia nigra pars reticulata, the major

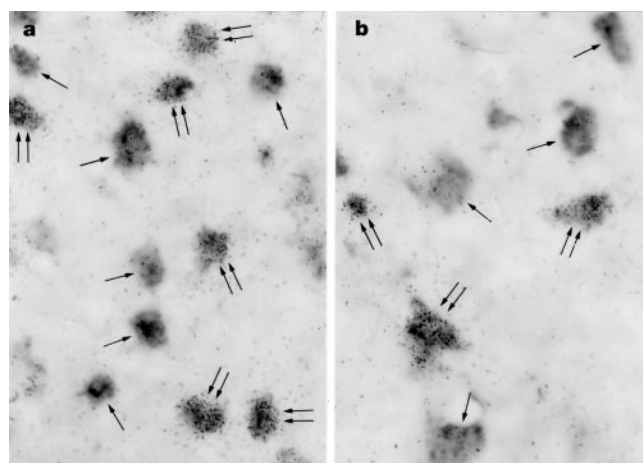


Figure 1 Emulsion autoradiogram illustrating co-expression of adenosine A_{2A} receptor mRNA (silver grains) and DARPP-32 mRNA (dark cells). Shown are subpopulations of medium spiny neurons in mouse (a) and rat (b) striatum. Single arrows indicate neurons that only express DARPP-32 mRNA. Double arrows indicate neurons that express both DARPP-32 mRNA and adenosine A_{2A} receptor mRNA.

output area of the basal ganglia, via the globus pallidus and subthalamic nucleus³. It is generally believed that activation of this 'indirect' striato-pallidal pathway reduces motor activity by inhibiting glutamatergic thalamo-cortical neurons³. Striato-pallidal neurons express high levels of the A_{2A} receptor subtype for the neuromodulator adenosine^{4,5}. Adenosine A_{2A} receptors are positively coupled to adenylyl cyclase via the olfactory-neuron-specific G protein, G_{olf} , and their activation stimulates the cAMP pathway^{6,7}. The psychostimulant caffeine is an adenosine receptor antagonist, able to produce a long-lasting increase in motor activity¹. Recent evidence obtained in A_{2A} receptor knockout mice confirms that the ability of caffeine to block adenosine A_{2A} receptors is responsible for the motor stimulant properties of this drug^{8,9}. However, almost nothing is known about the intracellular mechanisms underlying the behavioural effects of caffeine.

Striatal projection neurons express high levels of DARPP-32 (refs 2, 10). By double *in situ* hybridization, we found that messenger RNA for adenosine A_{2A} receptors was expressed in a subset of striatal neurons containing DARPP-32 (Fig. 1). The cells

expressing A_{2A} receptors represented about 50% of all DARPP-32-containing neurons. To determine whether DARPP-32 might be involved in mediating the behavioural effects of caffeine and selective adenosine A_{2A} receptor antagonists and agonists, we examined the effects of these drugs on motor activity in DARPP-32 knockout mice. Administration of 7.5 mg kg^{-1} of caffeine to wild-type mice produced, as expected, a large increase in motor activity, measured both as long-range movements (locomotion) (Fig. 2a, b) and as short-range movements (motility) (Fig. 2c), which were still present 100 min after drug administration (Fig. 2a). When the same dose of caffeine was injected into DARPP-32 knockout mice, a significant attenuation of the stimulant effect was observed on both locomotion (Fig. 2a, b) and motility (Fig. 2c). The DARPP-32-dependence of the stimulatory effect of caffeine on motor activity could be overcome by increasing the dose to 15 mg kg^{-1} (Fig. 2g).

To further evaluate the possible involvement of DARPP-32 in the behavioural response to blockade of adenosine A_{2A} receptors, we examined the motor stimulant effects produced by the selective A_{2A}

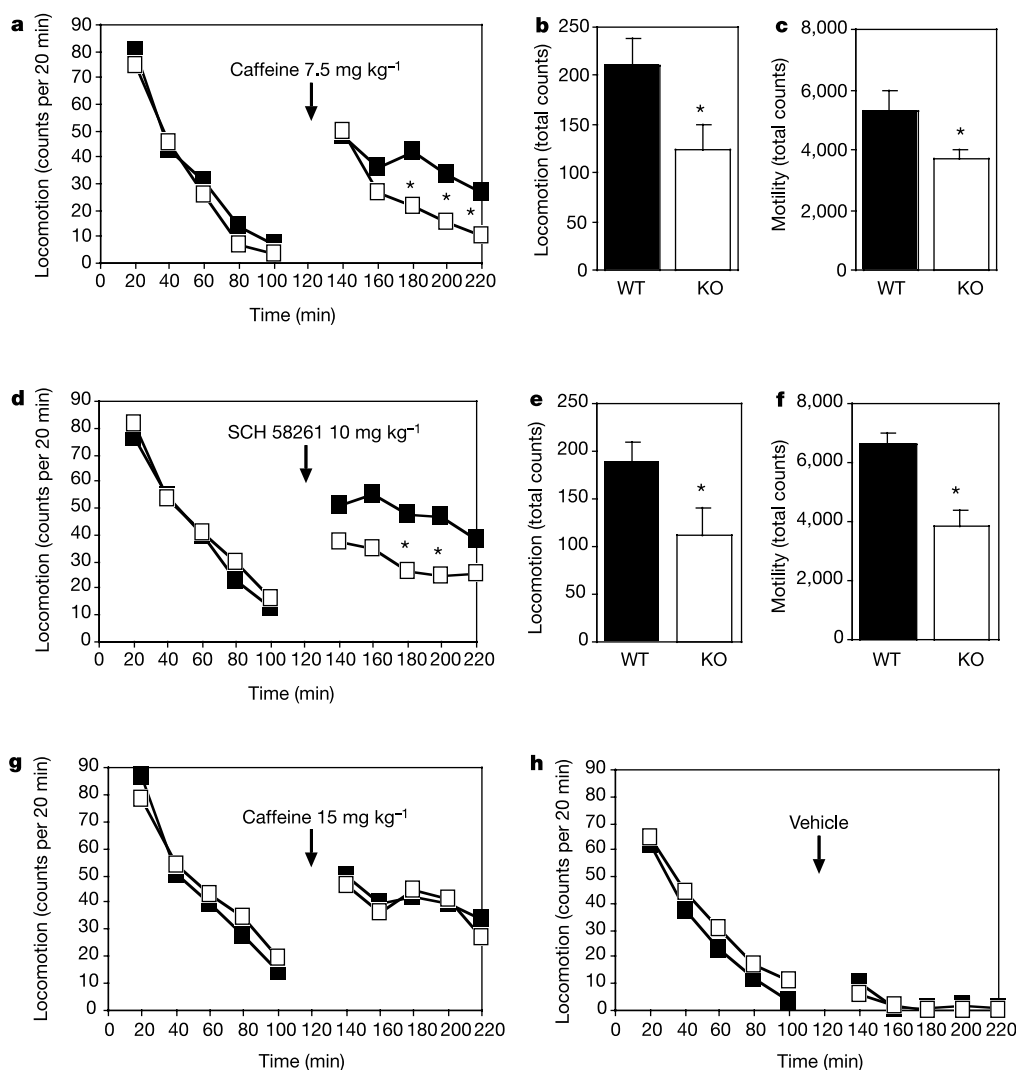


Figure 2 DARPP-32 is required for the stimulatory effects of caffeine and the selective adenosine A_{2A} receptor antagonist, SCH 58261, on motor activity. After a two-hour habituation period, wild-type (WT) mice (filled symbols) and DARPP-32 knockout (KO) mice (open symbols) received an injection of caffeine (7.5 or 15 mg kg^{-1} , i.p.) (a–c, g), SCH 58261 (10 mg kg^{-1} , i.p.) (d–f) or vehicle (h) and were immediately returned to their individual cages. a, d, g, h, Time course of the effects of caffeine, SCH 58261 or vehicle

on locomotion measured over 20-min intervals. b, e, Total locomotion counts determined during the 100-min period following administration of caffeine (b) or SCH 58261 (e). c, f, Total motility counts determined during 100 min following administration of caffeine (c) or SCH 58261 (f). Data represent means $\pm$ s.e.m. ($n = 15$ –20). Asterisk, $P < 0.05$ versus wild-type, Student's *t*-test.

receptor antagonist, SCH 58261, in wild-type and DARPP-32 knockout mice. As with caffeine we found that the increases in locomotion (Fig. 2d, e) and motility (Fig. 2f) induced by administration of SCH 58261 were reduced in DARPP-32 knockout mice.

Stimulation of adenosine A_{2A} receptors depresses motor activity and produces catalepsy¹¹. We found that, in wild-type mice, the A_{2A} receptor agonist CGS 21680 (0.1 mg kg^{-1} , administered intraperitoneally, i.p.) reduced motor activity by about 50% (Fig. 3, filled symbols). When the same dose of CGS 21680 was administered to DARPP-32 knockout mice, we observed a significant attenuation in the motor depressant effect of the A_{2A} receptor agonist (Fig. 3, open symbols).

The results shown in Figs 2 and 3 indicated that DARPP-32 is important in mediating the behavioural effects of caffeine and other adenosine A_{2A} receptor ligands. There is now compelling evidence indicating that DARPP-32 function is regulated by phosphorylation at multiple sites. Notably, our recent studies have indicated that Cdk5 phosphorylates DARPP-32 at Thr 75 and converts the protein into an inhibitor of PKA¹². Moreover, phospho-Thr 75 of DARPP-32 is dephosphorylated by protein phosphatase 2A (PP-2A), a process that contributes to signal amplification involving DARPP-32 (refs 13, 14). We therefore examined the effects of adenosine A_{2A} receptor antagonists on phosphorylation of DARPP-32 at Thr 75. Administration of caffeine (7.5 mg kg^{-1}) increased Thr 75 phosphorylation (Fig. 4a). SCH 58261 (10 mg kg^{-1}) mimicked the effect of caffeine on Thr 75 phosphorylation (Fig. 4b). Caffeine produced

an increase in Thr 75 phosphorylation, which was still significant 2 h after injection, whereas SCH 58261 was less effective at late time points (Fig. 4), probably owing to its shorter half-life¹⁵. The caffeine-induced increase in Thr 75 phosphorylation was abolished by the previous administration of roscovitine, an inhibitor of Cdk5 (data not shown), confirming that we were examining the Cdk5/PP-2A site of DARPP-32.

An increased phosphorylation at Thr 75 of DARPP-32 could be achieved by activation of Cdk5, by inhibition of PP-2A, or by both mechanisms. Cdk5 activity was measured in striatal extracts from mice which had been administered either 7.5 mg kg^{-1} of caffeine or vehicle. No difference was observed in the rate of Cdk5-catalysed substrate phosphorylation between caffeine- and vehicle-treated animals (Fig. 5a). To examine the role of PP-2A in mediating the effects of A_{2A} receptor regulation on Thr 75 phosphorylation, we used a slice preparation from mouse dorsal striatum. A decrease in phospho-Thr 75 was produced by the A_{2A} receptor agonist, CGS 21680. Moreover, this decrease was abolished by pre-incubation with okadaic acid, an inhibitor of PP-2A (Fig. 5b). Together, these studies suggest that adenosine A_{2A} receptors alter Thr 75 phosphorylation levels by regulating PP-2A activity.

These results clearly implicate DARPP-32 in the behavioural effects produced by caffeine and other adenosine A_{2A} receptor ligands. Furthermore, the data implicate phosphorylation at Thr 75 of DARPP-32 in the psychostimulant action of caffeine. It was recently demonstrated that phospho-Thr 75-DARPP-32 is an

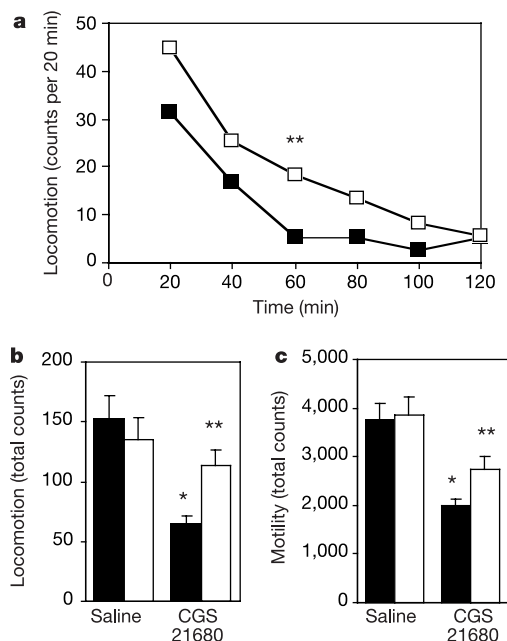


Figure 3 DARPP-32 is required for the depressant effect of the selective adenosine A_{2A} receptor agonist, CGS 21680, on motor activity. Wild-type mice (filled symbols) or DARPP-32 knockout mice (open symbols) were treated with vehicle or CGS 21680 (0.1 mg kg^{-1}) and placed in individual cages 10 min after the injection. The dose of CGS 21680 was chosen on the basis of previous studies indicating that it caused a substantial reduction in motor activity²³, without inducing catalepsy²⁴. **a**, Time course of the effect of CGS 21680 measured over 20-min intervals. **b**, **c**, Total locomotion counts (**b**) and total motility counts (**c**) determined during the 100-min period following administration of CGS 21680. Data represent means $\pm$ s.e.m. ($n = 20$ –25). Asterisk, $P < 0.001$ versus wild-type mice treated with vehicle; double asterisk, $P < 0.05$ versus wild-type mice treated with CGS 21680, Student's t -test. The effect of CGS 21680 was significantly reduced in DARPP-32 knockout mice ($P < 0.05$, two-way ANOVA).

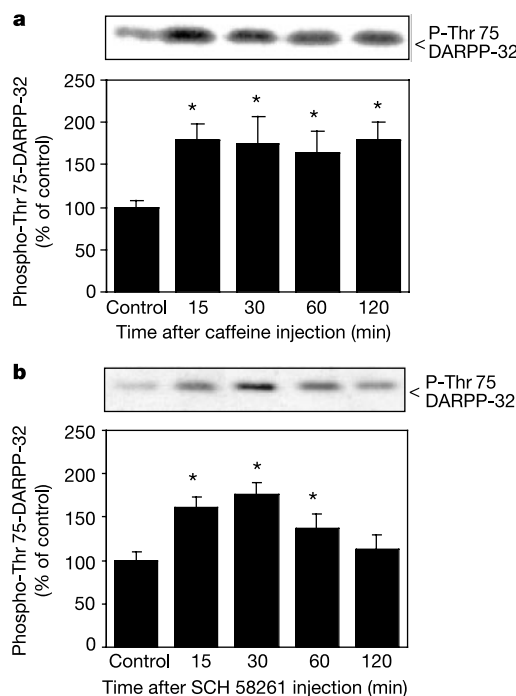


Figure 4 Effect of caffeine and SCH 58261 on DARPP-32 phosphorylation at Thr 75. Mice were treated i.p. with caffeine (7.5 mg kg^{-1}) (**a**) or SCH 58261 (10 mg kg^{-1}) (**b**) and decapitated at the indicated times after injection. The dose of caffeine was chosen based on studies indicating that 7.5 mg kg^{-1} of caffeine was able to produce a maximal increase in Thr 75 DARPP-32 phosphorylation (data not shown). The dose of SCH 58261 was chosen on the basis of previous studies indicating that 10 mg kg^{-1} of SCH 58261 was able to abolish the increase in DARPP-32 phosphorylation at Thr 34 induced by eticlopride, a dopamine D_2 receptor antagonist²⁵. The striatal levels of phospho-Thr 75-DARPP-32 were determined as described in Methods. Upper panels show representative autoradiograms; lower panels show summary of data expressed as means $\pm$ s.e.m. ($n = 10$). The amount of phosphorylated DARPP-32 is expressed as a percentage of that determined after vehicle administration. Asterisk, $P < 0.05$ versus vehicle injected mice, Student's t -test.

effective inhibitor of PKA¹². In addition, evidence was obtained suggesting that a cascade involving increased cAMP, activation of PKA and activation of PP-2A leads to the dephosphorylation of Thr 75 (ref. 14). The removal of the inhibitory constraint on PKA by dephosphorylation of Thr 75 provides a mechanism for amplification of the PKA signal transduction pathway¹⁴. Conversely, we hypothesize that the inhibition of the cAMP pathway produced by low caffeine levels would be amplified by the concomitant increase in phosphorylation at Thr 75 of DARPP-32 leading to further reduction of PKA activity. The present results indicate that a high (15 mg kg⁻¹) dose of caffeine is sufficient to shut down the PKA pathway without the need for this amplification. By the same logic, the similarity of the initial responses of wild-type and DARPP-

32 knockout mice to the motor stimulant effects of a lower caffeine concentration (7.5 mg kg⁻¹), in spite of the rapid increases in Thr 75 phosphorylation, can be explained by considering that, at this early stage, the level of caffeine is sufficiently high to produce near-maximal suppression of PKA activity. Nevertheless, as the levels of caffeine decrease, DARPP-32 and its phosphorylation at Thr 75 is used to maintain heightened locomotor responses.

In conclusion, we propose a model in which caffeine achieves its behavioural action in striato-pallidal neurons through a signalling cascade involving adenosine A_{2A} receptor blockade, decreased cAMP formation, decreased PKA activation, decreased PP-2A activation and increased phosphorylation of Thr 75 of DARPP-32. The caffeine-induced increase in the state of phosphorylation of Thr 75 would further lower PKA activity, thereby providing a positive-feedback amplification mechanism for shutting down adenosine A_{2A} receptor-stimulated PKA signalling pathways. □

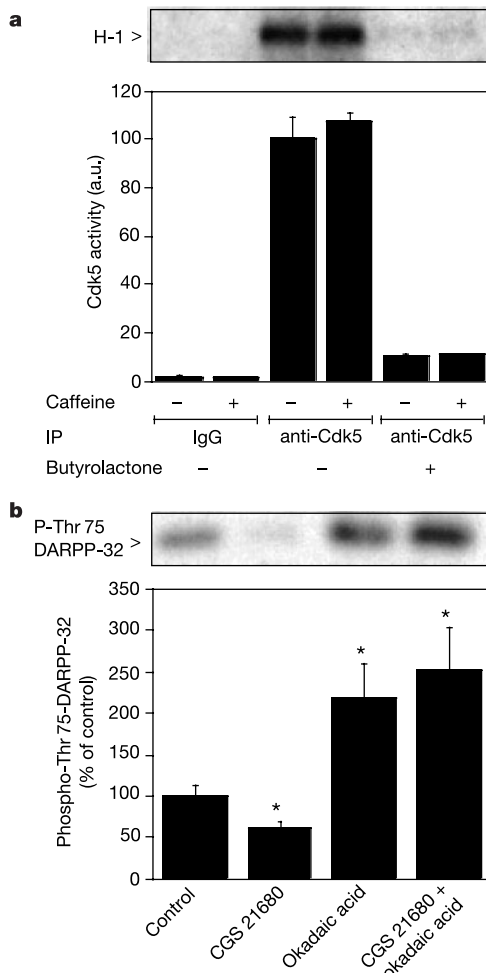


Figure 5 Adenosine A_{2A} receptor regulation of DARPP-32 phosphorylation at Thr 75 by modulation of PP-2A. **a**, Mice were treated with vehicle or 7.5 mg kg⁻¹ of caffeine and decapitated 15 min later. Dorsal striata were dissected out, homogenized in lysis buffer and subjected to immunoprecipitation (IP) with either a Cdk5 antibody (C8) or normal rabbit IgG. Cdk5 in the immunoprecipitates was assayed using histone H-1 as substrate in the absence or presence of 10 μ M butyrolactone, a specific Cdk5 inhibitor²². Upper panel, autoradiogram showing histone H-1 phosphorylation. Lower panel, data are means $\pm$ s.e.m. (n = 8). **b**, Slices from dorsal striatum were preincubated with adenosine deaminase for 30 min, to prevent the build up of endogenous adenosine and then incubated for 20 min in the presence of 1 μ M okadaic acid¹³ and for an additional 10 min in the presence of okadaic acid plus the A_{2A} receptor agonist CGS 21680 (1 μ M). Upper panel, autoradiogram showing phosphorylation at Thr 75 of DARPP-32. Lower panel, data are means $\pm$ s.e.m. (n = 7). The amount of phosphorylated DARPP-32 is expressed as a percentage of that determined after vehicle administration. Asterisk, P < 0.01 versus vehicle-treated group, Student's t -test.

Methods

In situ hybridization

Digoxigenin or ³⁵S-labelled riboprobes were prepared by *in vitro* transcription from complementary DNA clones corresponding to fragments of DARPP-32 (ref. 16) and adenosine A_{2A} receptor mRNA⁵, respectively. Coronal cryostat sections at different levels through striatum were hybridized as previously described¹⁷. Neurons containing a number of silver grains at least twofold higher than background were considered to be labelled for adenosine A_{2A} receptor mRNA.

Determination of adenosine A₁ receptor binding

The densities of adenosine A₁ receptors in wild-type and DARPP-32 knockout mice were determined by receptor autoradiography as previously described¹⁸, using [³H]1,3-dipropyl-8-cyclopentylxanthine ([³H]DPCPX; from DuPont-NEN) as the ligand.

Measurement of motor activity

We measured horizontal motor activity by means of an infrared-sensitive motion detector (Motron Products) consisting of 48 photosensors mounted in 4 $\times$ 4 cm squares and placed in the floor of the cage in two identical boxes. A motility count was defined as any movement covering two adjacent rows of photosensors and a locomotion count was defined as passing from one box to the other and crossing at least eight photosensors in the new box¹⁹. Wild type and DARPP-32 knockout mice²⁰ were generated from the offspring of DARPP-32^{+/+} $\times$ DARPP-32^{+/+} and DARPP-32^{-/-} $\times$ DARPP-32^{-/-} mating pairs. All mice were age-matched and only male offspring were used. No differences were observed between wild-type and DARPP-32 knockout mice in the densities of adenosine A_{2A} receptors¹⁸ or A₁ receptors ([³H]DPCPX binding in striatum and cortex of DARPP-32 knockout mice was 96.6 and 96.4% of wild type, respectively). The animals were accustomed to the experimental room for 40 min before the start of the experiments. To examine the motor stimulant effects of caffeine and SCH 58261, mice were first placed in individual cages for 2 h (habituation period). At the end of the habituation period the animals received an injection of caffeine (7.5 or 15 mg kg⁻¹, i.p.) or SCH 58261 (10 mg kg⁻¹, i.p.) and were immediately returned to their individual cages. To examine the motor depressant effect of CGS 21680, mice were treated with vehicle or 0.1 mg kg⁻¹ of CGS 21680 and placed in individual cages 10 min after the injection. To avoid the effects of decreased blood pressure caused by activation of peripheral A_{2A} receptors, mice were treated with the A_{2A} receptor antagonist 8-para-sulpho-theophylline (10 mg kg⁻¹, i.p.) 10 min before the administration of CGS 21680. 8-para-sulpho-theophylline does not cross the blood-brain barrier and, when administered alone, did not produce any significant biochemical or behavioural changes (data not shown).

Determination of phosphoThr75-DARPP-32

Male C57BL/6 mice (20–30 g; M&B) were injected intraperitoneally with vehicle or drugs and killed by decapitation. When combinations of two drugs were used, the mice were killed 15 min after the second injection. The heads of the animals were immediately immersed in liquid nitrogen for six seconds. The brains were then removed and the dorsal parts of the striata rapidly (20 s) dissected out on an ice-cold surface, sonicated in 750 μ l of 1% sodium dodecylsulphate and boiled for 10 min. For slice experiments, we prepared brain coronal slices (200 μ m) using a vibroslice (Campden Instruments). Dorsal striata were then dissected out from each slice under a microscope. Two striatal slices were placed in individual 5-ml polypropylene tubes containing 2 ml of Krebs–Ringer bicarbonate buffer (KRB): 118 mM NaCl, 4.7 mM KCl, 1.3 mM CaCl₂, 1.5 mM MgSO₄, 1.2 mM KH₂PO₄, 25 mM NaHCO₃, 11.7 mM glucose, equilibrated with 95% O₂/5% CO₂ (v/v), pH 7.3. The samples were equilibrated at 30 $^{\circ}$ C for two 30-min periods, each followed by replacement of the medium with 2 ml of fresh KRB. Test substances were then added. After incubation, the solutions were rapidly removed, the slices were sonicated in 200 μ l of 1% sodium dodecylsulphate and boiled for 10 min. Phospho-Thr 75-DARPP-32 was detected as described using a polyclonal antibody¹². In some experiments, a monoclonal antibody (C24-5a) (diluted 1:10,000)²¹ generated against DARPP-32, which is not phosphorylation-state-specific, was used to measure the total amount of DARPP-32. Antibody binding was revealed using goat anti-mouse horseradish peroxidase (HRP)-linked IgG (diluted 1:10,000), and the enhanced chemiluminescence (ECL)

immunoblotting detection system. Chemiluminescence was detected by autoradiography. Quantification of the phospho-DARPP-32 bands was done by densitometry, using NIH Image (version 1.52) software.

Determination of Cdk5 activity

Mice were treated intraperitoneally with vehicle or 7.5 mg kg⁻¹ of caffeine and decapitated 15 min later. The brains were rapidly removed, and the dorsal striata were dissected out and homogenized in lysis buffer. Cdk5 was immunoprecipitated with anti-Cdk5 (C8) antibody and its activity was assayed in the absence or presence of butyrolactone (10 μM) by using histone H-1 as the substrate as described²².

Received 21 February; accepted 3 April 2002; doi:10.1038/nature00817.

1. Fredholm, B. B., Battig, K., Holmen, J., Nehlig, A. & Zvartau, E. E. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacol. Rev.* **51**, 83–133 (1999).
2. Walaas, S. L., Aswad, D. W. & Greengard, P. DARPP-32, a dopamine- and cAMP-regulated phosphoprotein enriched in dopamine-innervated brain regions. *Nature* **301**, 69–71 (1983).
3. Gerfen, C. R. The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu. Rev. Neurosci.* **15**, 285–320 (1992).
4. Schiffman, S. N. & Vanderhaeghen, J.-J. Adenosine A2 receptors regulate the gene expression of striatopallidal and striatonigral neurons. *J. Neurosci.* **13**, 1080–1087 (1993).
5. Fink, J. S. *et al.* Molecular cloning of the rat A2 adenosine receptor: selective co-expression with D2 dopamine receptor in rat striatum. *Mol. Brain Res.* **14**, 186–195 (1992).
6. Kull, B., Svenningsson, P. & Fredholm, B. B. Adenosine A2A receptors are colocalized with and activate Golf in rat striatum. *Mol. Pharmacol.* **58**, 771–777 (2000).
7. Fredholm, B. B. Activation of adenylate cyclase from rat striatum and tuberculum olfactorium by adenosine. *Med. Biol.* **55**, 262–267 (1977).
8. Ledent, C. *et al.* Aggressiveness, hypoalgesia and increased blood pressure in mice deficient for the adenosine A2a receptor. *Nature* **388**, 674–678 (1997).
9. El Yacoubi, M. *et al.* The stimulant effects of caffeine on locomotor behaviour in mice are mediated through its blockade of adenosine A(2A) receptors. *Br. J. Pharmacol.* **129**, 1465–1473 (2000).
10. Ouimet, C. C., Langley-Guillion, K.-C. & Greengard, P. Quantitative immunohistochemistry of DARPP-32-expressing neurons in the rat caudateputamen. *Brain Res.* **808**, 8–12 (1998).
11. Snyder, S. H., Katims, J. J., Annau, Z., Bruns, R. F. & Daly, J. W. Adenosine receptors and the behavioural actions of methylxanthines. *Proc. Natl Acad. Sci. USA* **78**, 3260–3264 (1981).
12. Bibb, J. *et al.* Phosphorylation of DARPP-32 by Cdk5 modulates dopamine signalling in neurons. *Nature* **402**, 669–671 (1999).
13. Nishi, A., Snyder, G. L., Nairn, A. C. & Greengard, P. Role of calcineurin and protein phosphatase-2A in the regulation of DARPP-32 dephosphorylation in neostriatal neurons. *J. Neurochem.* **72**, 2015–2021 (1999).
14. Nishi, A. *et al.* Amplification of dopaminergic signalling by a positive feedback loop. *Proc. Natl Acad. Sci. USA* **97**, 12840–12845 (2000).
15. Ongini, E. SCH 58261: a selective A2A adenosine receptor antagonist. *Drug Dev. Res.* **42**, 63–70 (1997).
16. Ehrlich, M. E., Kurihara, T. & Greengard, P. Rat DARPP-32: cloning, sequencing, and characterization of the cDNA. *J. Mol. Neurosci.* **2**, 1–10 (1990).
17. Le Moine, C. & Bloch, B. D1 and D2 dopamine receptors gene expression in the rat striatum: sensitive cRNA probes demonstrate prominent segregation of D1 and D2 mRNA in distinct neuronal population of the dorsal and ventral striatum. *J. Comp. Neurol.* **355**, 418–427 (1995).
18. Svenningsson, P. *et al.* Dopamine D1 receptor-induced gene transcription is modulated by DARPP-32. *J. Neurochem.* **75**, 248–257 (2000).
19. Ögren, S.-Ö., Köhler, C., Fuxe, K. & Ångeby, K. In *Dopaminergic Ergot Derivatives and Motor Function* (eds Fuxe, K. & Calne, D. B.) 225–266 (Pergamon, Oxford, 1979).
20. Fienberg, A. A. *et al.* DARPP-32: regulator of the efficacy of dopaminergic neurotransmission. *Science* **281**, 838–842 (1998).
21. Hemmings, H. C. J. & Greengard, P. DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein: regional, tissue, and phylogenetic distribution. *J. Neurosci.* **6**, 1469–1481 (1986).
22. Liu, F. *et al.* Regulation of cyclin-dependent kinase 5 and casein kinase 1 by metabotropic glutamate receptors. *Proc. Natl Acad. Sci. USA* **98**, 11062–11068 (2001).
23. Rimondini, R., Ferré, S., Giménez-Llort, L., Ögren, S.-Ö. & Fuxe, K. Differential effects of selective adenosine A1 and A2 receptor agonists on dopamine receptor agonist-induced behavioural responses in rats. *Eur. J. Pharmacol.* **347**, 153–158 (1998).
24. Kafka, S. H. & Corbett, R. Selective adenosine A2A receptor/dopamine D2 receptor interactions in animal models of schizophrenia. *Eur. J. Pharmacol.* **295**, 147–154 (1996).
25. Svenningsson, P. *et al.* Regulation of the phosphorylation of the dopamine- and cAMP-regulated phosphoprotein of 32 kDa *in vivo* by dopamine D1, dopamine D2, and adenosine A2A receptors. *Proc. Natl Acad. Sci. USA* **97**, 1856–1860 (2000).

Acknowledgements

We thank R. Rimondini and S.-Ö. Ögren for help in performing behavioural experiments and A. Nishi for discussions. This work was supported by the Swedish Research Council (G.F. and B.B.F.), the Swedish Society for Medical Research (M.L.), the Foundation Blanceflor Boncompagni-Ludovisi, née Bildt (L.P.) and by funding from the National Institute of Mental Health and the National Institute of Drug Abuse (A.C.N. and P.G.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.F. (e-mail: gilberto.fisone@neuro.ki.se).

Sperm from neonatal mammalian testes grafted in mice

Ali Honaramooz*, Amy Snedaker*, Michele Boiani†, Hans Schöler†, Ina Dobrinski* & Stefan Schlatt†‡

* Male Germ Cell Biology Group, and † Germline Development Group, Center for Animal Transgenesis and Germ Cell Research, New Bolton Center, School of Veterinary Medicine, University of Pennsylvania, Kennett Square, Pennsylvania 19348, USA

‡ Institute of Reproductive Medicine, University Münster, Domagkstrasse 11, D-48149 Münster, Germany

Spermatogenesis is a productive and highly organized process that generates virtually unlimited numbers of sperm during adulthood. Continuous proliferation and differentiation of germ cells occur in a delicate balance with other testicular compartments, especially the supporting Sertoli cells¹. Many complex aspects of testis function in humans and large animals have remained elusive because of a lack of suitable *in vitro* or *in vivo* models. Germ cell transplantation has produced complete donor-derived spermatogenesis in rodents^{2–6} but not in other mammalian species^{7–9}. Production of sperm in grafted tissue from immature mammalian testes and across species has not yet been accomplished. Here we report the establishment of complete spermatogenesis by grafting testis tissue from newborn mice, pigs or goats into mouse hosts. This approach maintains structural integrity and provides the accessibility that is essential for studying and manipulating the function of testes and for preserving the male germ line. Our results indicate that this approach is applicable to diverse mammalian species.

Transplantation of spermatogonial stem cells from fertile donor mice to the testes of infertile recipient mice results in complete spermatogenesis^{2–3,10} and autologous transplantation is successful in the monkey¹¹, which opens up the field for medical applications¹². Cross-species transplantation of spermatogonial stem cells from donor rats or hamsters to recipient mice results in the establishment of rat or hamster spermatogenesis in the mouse testis^{5,6}; however, the transplantation of germ cells from phylogenetically more distant species including rabbits, dogs, pigs, bulls, horses and primates into

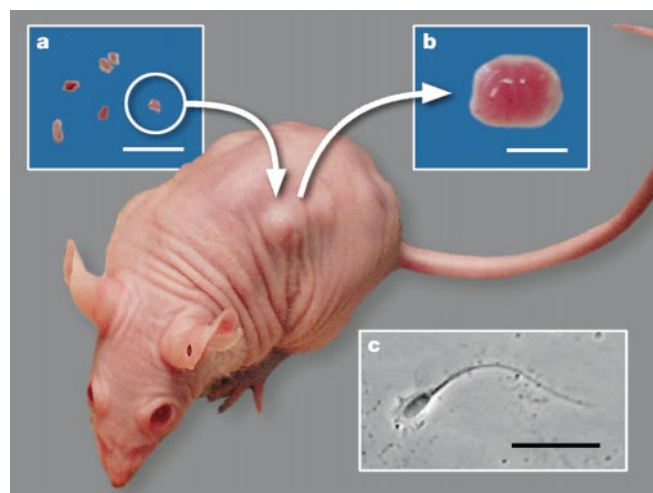


Figure 1 Grafting of testis tissue from newborn piglets under the skin of nude mice. **a, b**, The size of the pig testis grafts at the time of transplantation was about 0.5–1 mm in diameter (**a**) and expanded to 4–8 mm at 10 weeks after grafting (**b**). Scale bars, 5 mm. Most grafts from later time points contained sperm. **c**, A sperm extracted from a week 27 graft. Scale bar, 20 μm.

mouse testes, does not result in spermatogenesis beyond the stage of spermatogonial proliferation⁷⁻⁹, probably owing to the incompatibility of microenvironments. We therefore developed a method of grafting small pieces of testis tissue into mouse hosts as an alternate approach to maintaining and propagating male germ cells that could be applied more readily to different mammalian species.

Fragments of testicular tissue (0.5–1 mm³) from neonatal mice, pigs or goats were grafted under the back skin of castrated immunodeficient mice. Groups of recipient mice were killed after grafting at intervals of 2–4 weeks, and the status of graft survival and spermatogenesis were assessed. Overall, more than 60% of the neonatal testicular grafts from all three donor species survived. All recovered grafts had increased in volume, some more than 100-fold, and ultimately produced mature sperm (Fig. 1). The recipient mice ($n = 94$), some kept as long as a year, did not show signs of illness or evidence of neoplasia in any of the recovered grafts ($n = 477$).

Histological analysis showed that there was full differentiation of the mouse testis allografts. Complete spermatogenesis was observed in the grafts of newborn murine testis, which originally contain primitive gonocytes as the only type of germ cell (ref. 13 and Fig. 2). The kinetics of spermatogenic progression were identical to that in intact mouse testis¹⁴: spermatocytes were the most advanced germ cells observed after 2 weeks, and the first round of spermatogenesis was complete 4 weeks after grafting. In many seminiferous tubules, however, there was a dilation of the lumen accompanied by a disorganized epithelium and premature sloughing of postmeiotic germ cells, probably as a result of the block of fluid flow^{15,16}.

Xenografting of newborn pig testis fragments into mice resulted in complete spermatogenesis that followed the developmental patterns of pig testis. There was a gradual progression from the primitive seminiferous cords (Fig. 3a) into fully developed seminiferous tubules with all stages of spermatogenic cells present (Fig. 3b, c). Thereby, in contrast to xenotransplantation of isolated testicular cells⁷⁻⁹, xenografting of pig testis fragments led to the production of sperm. Spermatogenesis developed earlier in grafted tissue than in the pig testis. Round spermatids were first present at 12 weeks after grafting, whereas they appear at or after 14 weeks of age in pig testes. In addition, the diameter of seminiferous tubules

increased faster in grafts; at week 8 it was about 120 μm (range 72–187 μm), as compared with about 56 μm (range 47–67 μm) in intact pig testes. Some tubules showed asynchronous development (Fig. 3b), which was not observed in mouse-to-mouse grafts. In contrast to mouse grafts, pig grafts examined at much later time points (up to a year) showed mostly complete, morphologically normal spermatogenesis, with the number of sperm per gram (for example, 112×10^6 sperm isolated from a 0.75-g graft) comparable to that in intact pig testes¹⁷⁻¹⁹.

Goat testis xenografts also developed from an immature state (Fig. 3e) to full spermatogenesis with seminiferous tubules showing normal stages of goat germ cell development (Fig. 3f). Mature and viable sperm could also be isolated from the grafts in high concentrations (for example, 67×10^6 sperm per g).

Sperm recovered from testis grafts of all three species into mice were viable and functional, as shown by fertilization after intracytoplasmic injection into mouse oocytes (Fig. 4). Homologous mouse embryos produced using this system showed normal fetal development when they were transferred into pseudopregnant mice. The production of mature and functionally competent sperm from three phylogenetically different species indicates that testis tissue grafting will be applicable to many different mammalian species.

We showed that spermatogenesis is initiated and completed in fresh testicular tissue grafted into mice. To extend this powerful approach to situations in which grafting of freshly collected tissue is not possible, we also investigated the cooling or cryopreservation of pig testis fragments before grafting. Indeed, testis tissue preserved by either refrigeration for up to 2 d or by cryopreservation for longer periods maintained the potential to develop complete spermatogenesis and steroidogenesis when grafted into mice (Fig. 3d). This observation makes the technique immediately applicable to a clinical or field situation.

Normal spermatogenesis requires the combined effects of gonadotrophins (follicle-stimulating hormone, FSH, and leutinizing hormone, LH), androgens and possibly other hormones. The endocrine function of the testis is controlled by a feedback loop known as the hypothalamic–pituitary–gonadal axis. Testosterone released from Leydig cells and inhibin secreted from Sertoli cells

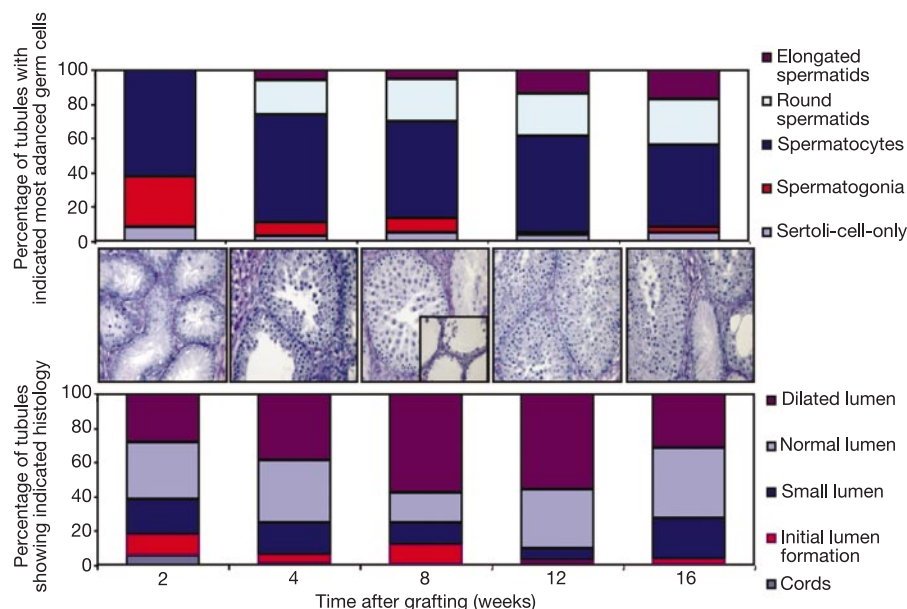


Figure 2 Histological evaluation of the developmental potential of newborn mouse testis tissue grafted under the back skin of nude mice. Top, progression of spermatogenesis illustrated by the relative abundance of seminiferous tubules containing the indicated

germ cell as the most advanced stage. Middle, representative micrographs of the testicular grafts. Bottom, progression of immature cords to fully differentiated seminiferous tubules showing the various degrees of lumen formation.

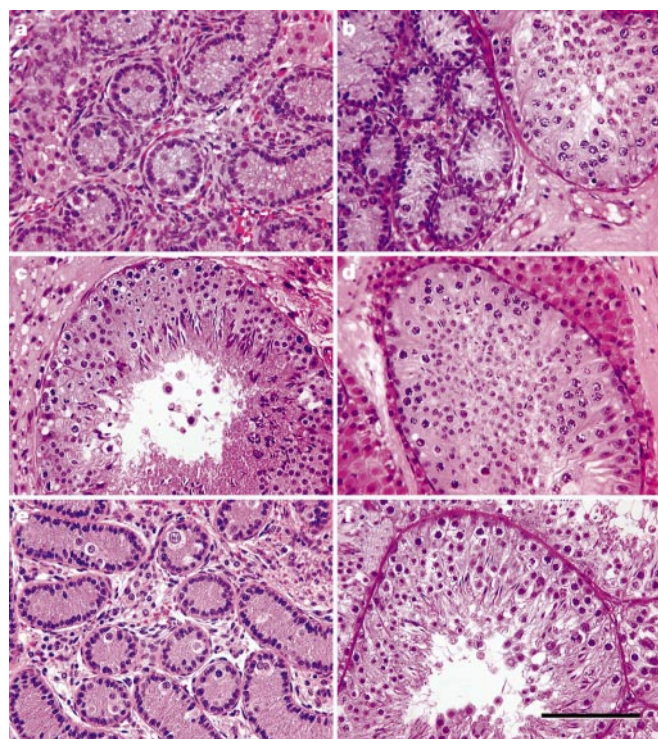


Figure 3 Histological appearance of pig or goat testis tissue before and after grafting into mice. **a–c**, Neonatal donor pig testis at the time of grafting (**a**), and grafted tissue at 12 weeks (**b**) or 18 weeks (**c**). Note the changes from seminiferous cords (**a**) to fully developed seminiferous tubules with complete spermatogenesis (**c**). Asynchronous tubular development is apparent in a week 12 graft (**b**). **d**, Complete spermatogenic progression in a week 16 pig graft that was cryopreserved before grafting. **e, f**, One-month-old donor goat testis at the time of grafting (**e**) and grafted tissue at week 10 (**f**). Scale bar, 100 μ m.

negatively control the release of FSH and LH from the pituitary. In castrated animals and in the absence of testosterone, serum quantities of FSH and LH increase markedly. Our results from both allografts and xenografts indicate that the feedback loop that controls the release of FSH is initiated in grafted mice, as the serum quantities of FSH were significantly lower than in castrated mice and remained intermediate between the intact and castrate range. In mouse-to-mouse grafts, increasing damage to the seminiferous epithelium from fluid stasis may diminish feedback from the grafted tissue and cause concentrations of serum FSH to reach castrate concentrations after week 8. This was not observed in recipients of xenografts. A stable and functional feedback loop existed between the pituitary of the host and the Leydig cells in the allografts or xenografts in all recipient mice, as evident by a significant increase in the weight of the seminal vesicles (which are highly androgen-dependent and thereby a reliable bioassay for androgens) and by the significantly increased amounts of serum testosterone as compared with castrated control animals. This indicates that the grafts fully supplement androgens to the castrated recipients.

Although grafting of testis tissue was developed as a tool for androgen substitution in the 1950s (ref. 20) and has been subsequently applied to study steroidogenesis and spermatogenesis^{21,22}, to our knowledge this is the first report of the initiation and maintenance of full spermatogenesis and steroidogenesis in testis tissue from different donor species grafted in mice. The completion of spermatogenesis in testis xenografts originating from neonatal pigs and goats indicates that it is feasible to initiate and maintain spermatogenesis in a phylogenetically distant species. Murine gonadotrophins effectively support the development, differentiation

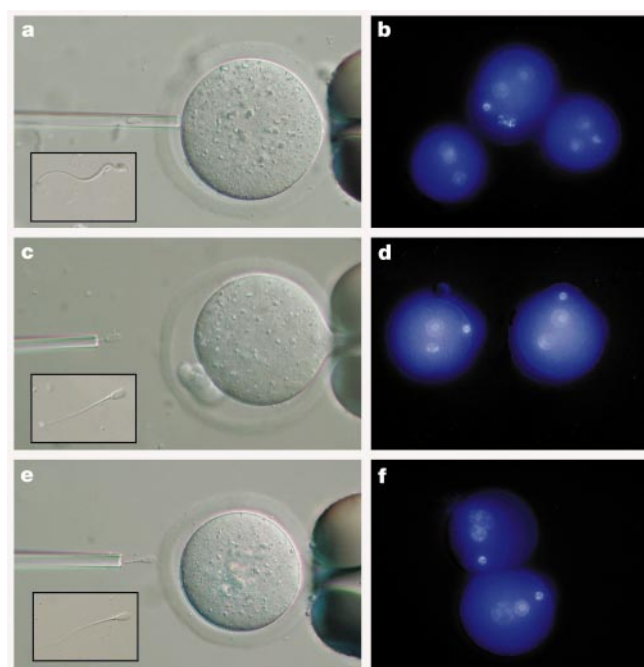


Figure 4 Sperm morphology and oocyte-activating competence of sperm from grafts of neonatal testes. Testicular sperm were isolated from mouse-to-mouse (**a**), pig-to-mouse (**c**) or goat-to-mouse (**e**) testis grafts before intracytoplasmic sperm injection into mouse oocytes. After injection, resumption of meiosis, polar body extrusion and pronuclear formation can be seen to occur within 6 h. Mouse oocytes injected with sperm from any of the three species (mouse, **b**; pig, **d**; goat, **f**) show two pronuclei and a second polar body (Hoechst 33342 staining).

and maintenance of these testicular xenografts despite species specificity, and exogenous hormone supplementation is not needed.

There are several immediate, highly relevant applications for testis tissue grafting. First, testis grafting provides a new option for male germline preservation. In contrast to the conventionally used method of sperm cryopreservation, this procedure provides a potentially inexhaustible source of male gametes, even from immature gonads. Unlike autologous transplantation of isolated germ cells to restore fertility in an individual after cancer therapy, xenologous grafting and use of the resultant sperm for assisted fertilization will eliminate the potential risk of transmitting tumour cells. Second, grafting fresh or preserved testis tissue offers an invaluable tool for conserving endangered species or valuable livestock by allowing sperm production from immature males. Third, the accessibility of the tissue in the mouse host makes it possible to manipulate spermatogenesis and steroidogenesis in a controlled manner that is not feasible in the donor animal and certainly not feasible in humans. This in turn will allow analysis of the effects of toxicants and potential male contraceptives on testis function in the target species. Last, grafting of testis tissue from experimental animal strains will provide geneticists with a previously unavailable tool to study germ cell development and even to produce gametes from animals with poor viability, such as neonatal lethal transgenic, mutant or cloned animals. □

Methods

Donor testis tissue

Mouse donor testes were obtained from ICR or B6C3 F₁ (C57BL/6 $\times$ C3H/He) pups aged 1–2 d. Testes were cut in half and each half served as a graft. Pig and goat testes were obtained from routine castration of animals aged 1 or 4 weeks, respectively. We prepared small fragments (~ 0.5 – 1 mm³) of these testis tissues. Before being grafted into recipient mice, all testis fragments were maintained in DMEM on ice. Comparable tissue pieces were fixed at the time of grafting as a reference for graft development.

Experimental surgery

Six-week-old male immunodeficient NCr mice (Taconic) were anaesthetized, castrated, and received eight transverse linear incisions (about 0.5–1 cm in length) into the dorsal skin. A fragment of the testis tissue was affixed subcutaneously near the incision with braided silk surgical suture, and the incisions were closed with wound clips (Michel Clips 7.5 mm, Miltex). As a control, we carried out sham-operations on groups of intact and castrated mice.

Sample processing and microscopy

At the time of death, the mice were anaesthetized and bled by heart puncture. Serum was kept frozen until analysis for hormone concentrations. We fixed surviving grafts in Bouin's solution and stored them in 70% ethanol until processing for histology. The weight of seminal vesicles was documented. We carried out the following histological analyses: photographic documentation of the histology, measurement of seminiferous tubule diameter, evaluation of the relative number of seminiferous tubules showing differentiation, and determination of the prevalence of the most advanced type of germ cell.

Isolation and storage of sperm from excised grafts

To obtain sperm from excised testis grafts, the tissue was minced and dispersed in Whittingham medium at 4 °C (ref. 23). The medium was supplemented with bovine serum albumin (BSA, 3% w/v)²⁴. The crude cell suspension was either used fresh or aliquoted in cryovials, snap-frozen in liquid nitrogen and maintained at –196 °C.

Intracytoplasmic injection of sperm into mouse oocytes

Metaphase II mouse oocytes and sperm from mouse, pig and goat grafts were micromanipulated in HEPES-buffered CZB medium (with glucose, BSA-free, polyvinylpyrrolidone 1% w/v)²⁵ using a piezo actuator (Prime-Tech) and DIC optics (Nikon) at 28 °C. Briefly, individual sperm were aspirated into a 10-µm (inner diameter) blunt-end borosilicate microcapillary loaded with mercury. For mouse sperm, the application of 1–5 piezo pulses to the neck region, held at the tip of the capillary, allowed partition of the sperm head from the tail. The sperm or sperm head was immediately injected deep into the oocyte along with a minimal amount of carrier medium. The injected oocytes were allowed to recover in the micromanipulation drop on the stage of the microscope for 15 min, transferred to culture medium (M16) and incubated at 37 °C under 5% CO₂ in air. As an indication of the sperm functional competence, we monitored the extrusion of the oocyte's second polar body and formation of the sperm-derived pronucleus by DNA staining with Hoechst 33342 at 2 and 8 h after intracytoplasmic injection²⁶.

Storage and cryopreservation of testis tissue

Testes of piglets were either maintained at 4 °C in saline solution for 2 d or were cut into small fragments, immersed in fetal bovine serum, DMEM and dimethyl sulphoxide at a 1:3:1 ratio, and cooled in an alcohol bath at a rate of –1 °C min^{–1} to –70 °C, before storing at –196 °C (refs 6, 7).

Hormone analysis

We measured serum amounts of testosterone by heterologous radioimmunoassay after double extraction²⁷. Intra- and interassay variations were 5.0 and 8.2%, respectively. Circulating concentrations of FSH were determined by a commercially available system (Rat FSH Assay System, Amersham Pharmacia), according to the manufacturer's instructions without magnetic separation. Intra- and interassay variations were 5.6 and 5.1%, respectively.

Received 7 March; accepted 23 May 2002; doi:10.1038/nature00918.

- Russell, L. D., Ettlin, R. A., SinhaHikim, A. P. & Clegg, E. D. in *Histological and Histopathological Evaluation of the Testis* (eds Russell, L. D., Ettlin, R. A., SinhaHikim, A. P. & Clegg, E. D.) 1–40 (Cache River, Clearwater, 1990).
- Brinster, R. L. & Zimmerman, J. W. Spermatogenesis following male germ cell transplantation. *Proc. Natl Acad. Sci. USA* **91**, 11298–11302 (1994).
- Brinster, R. L. & Avarbock, M. R. Germline transmission of donor haplotype following spermatogonial transfer. *Proc. Natl Acad. Sci. USA* **91**, 11303–11307 (1994).
- Jiang, F. X. & Short, R. V. Male germ cell transplantation in rats: Apparent synchronization of spermatogenesis between host and donor seminiferous epithelia. *Int. J. Androl.* **18**, 326–330 (1995).
- Clouthier, D. E., Avarbock, M. R., Maika, S. D., Hammer, R. E. & Brinster, R. L. Rat spermatogenesis in mouse testis. *Nature* **381**, 418–421 (1996).
- Ogawa, T., Dobrinski, I., Avarbock, M. R. & Brinster, R. L. Xenogeneic spermatogenesis following transplantation of hamster germ cells to mouse testes. *Biol. Reprod.* **60**, 515–521 (1999).
- Dobrinski, I., Avarbock, M. R. & Brinster, R. L. Transplantation of germ cells from rabbits and dogs into mouse testes. *Biol. Reprod.* **61**, 1331–1339 (1999).
- Dobrinski, I., Avarbock, M. R. & Brinster, R. L. Germ cell transplantation from large domestic animals into mouse testes. *Mol. Reprod. Develop.* **57**, 270–279 (2000).
- Nagano, M., McCarrey, J. R. & Brinster, R. L. Primate spermatogonial stem cells colonize mouse testis. *Biol. Reprod.* **64**, 1409–1416 (2001).
- Ogawa, T., Dobrinski, I., Avarbock, M. R. & Brinster, R. L. Transplantation of male germ line stem cells restores fertility in infertile mice. *Nature Med.* **6**, 29–34 (2000).
- Schlatt, S. *et al.* Germ cell transfer into rat, bovine, monkey and human testes. *Hum. Reprod.* **14**, 144–150 (1998).
- Schlatt, S., von Schönfeldt, V. & Schepers, A. G. Male germ cell transplantation: An experimental approach with a clinical perspective. *Brit. Med. Bull.* **56**, 577–587 (2000).
- Orth, J. M. in *Cell and Molecular Biology of the Testis* (eds Desjardins, C. & Ewing, L. L.) 3–42 (Oxford Univ. Press, New York, 1993).
- McCarrey, J. R. in *Cell and Molecular Biology of the Testis* (eds Desjardins, C. & Ewing, L. L.) 58–89 (Oxford Univ. Press, New York, 1993).

- Eddy, E. M. *et al.* Targeted disruption of the estrogen receptor gene in male mice causes alteration of spermatogenesis and infertility. *Endocrinology* **137**, 4796–4805 (1996).
- Lee, K.-H. *et al.* Estrogen receptor has a functional role in the mouse rete testis and efferent ductules. *Biol. Reprod.* **63**, 1873–1880 (2000).
- Amann, R. P. in *The Testis* (eds Johnson, A. D., Gomes, W. R. & Vademarck, N. L.) 433–482 (Academic, New York, 1970).
- Swierstra, E. E. A comparison of spermatozoa production and spermatozoa output of Yorkshire and Lacombe boars. *J. Reprod. Fertil.* **17**, 459–469 (1968).
- Swierstra, E. E. Sperm production of boars measured from epididymal sperm reserves and quantitative testicular histology. *J. Reprod. Fertil.* **27**, 91–99 (1971).
- Deanesly, R. Spermatogenesis and endocrine activity in grafts of frozen and thawed rat testis. *J. Endocrinol.* **11**, 201–206 (1954).
- Kuopio, T., Savouras, P. O., Pelliniemi, L. J. & Huhtaniemi, I. T. Transplantation of newborn rat testis under the kidney capsule of adult host as a model to study the structure and function of Leydig cells. *J. Androl.* **10**, 335–345 (1989).
- Johnson, L., Suggs, L. C., Norton, Y. M. & Zeh, W. C. Effect of developmental age or time after transplantation on Sertoli cell number and testicular size in inbred Fischer rats. *Biol. Reprod.* **54**, 948–959 (1996).
- Fraser, L. R. & Drury, L. M. The relationship between sperm concentration and fertilization *in vitro* of mouse eggs. *Biol. Reprod.* **13**, 513–518 (1975).
- Wakayama, T., Whittingham, D. G. & Yanagimachi, R. Production of normal offspring from mouse oocytes injected with spermatozoa cryopreserved with or without cryoprotection. *J. Reprod. Fertil.* **112**, 11–17 (1998).
- Chatot, C. L., Lewis, J. L., Torres, I. & Ziomek, C. A. Development of 1-cell embryos from different strains of mice in CZB medium. *Biol. Reprod.* **42**, 432–440 (1990).
- Goud, P. T., Goud, A. P., Rybouchkin, A. V., De Sutter, P. & Dhont, M. Chromatin decondensation, pronucleus formation, metaphase entry and chromosome complements of human spermatozoa after intracytoplasmic sperm injection into hamster oocytes. *Hum. Reprod.* **13**, 1336–1345 (1998).
- Chandolia, R. K., Weinbauer, G. F., Simoni, M., Behre, H. M. & Nieschlag, E. Comparative effects of chronic administration of the non-steroidal antiandrogens flutamide and Casodex on the reproductive system of the adult male rat. *Acta Endocrinol.* **125**, 547–555 (1991).

Acknowledgements

We thank R. Gosden who stimulated the development of the approach; R. Sandhowe-Klawerkamp and I. Upmann for technical assistance; and J. Hayden for preparing the illustrations. This study was supported by the NIH, United States Department of Agriculture/Cooperative State Research, Education and Extension Service/National Research Initiative Competitive Grants Program (I.D.), by the Marion Dilley and David George Jones Funds and the Commonwealth and General Assembly of Pennsylvania (M.B., S.S. and H.S.), an award from the Merck Foundation (A.S.) and by a Heisenberg fellowship from the Deutsche Forschungsgemeinschaft (S.S.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to I.D. (e-mail: dobrinski@vet.upenn.edu).

Distalization of the *Drosophila* leg by graded EGF-receptor activity

Gerard Campbell

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Arthropods and higher vertebrates both possess appendages, but these are morphologically distinct and the molecular mechanisms regulating patterning along their proximodistal axis (base to tip) are thought to be quite different. In *Drosophila*, gene expression along this axis is thought to be controlled primarily by a combination of transforming growth factor-β (TGF-β) and Wnt signalling from sources of ligands, Decapentaplegic (Dpp) and Wingless (Wg), in dorsal and ventral stripes, respectively^{1–3}. In vertebrates, however, proximodistal patterning is regulated by receptor tyrosine kinase (RTK) activity from a source of ligands, fibroblast growth factors (FGFs), at the tip of the limb bud⁴. Here I revise our understanding of limb development in flies and show that the distal region is actually patterned by a distal-to-proximal gradient of RTK activity, established by a source of epidermal

growth factor (EGF)-related ligands at the presumptive tip. This similarity between proximodistal patterning in vertebrates and flies supports previous suggestions^{5,6} of an evolutionary relationship between appendages/body-wall outgrowths in animals.

Wingless (Wg) and Decapentaplegic (Dpp) regulate proximodistal patterning in the *Drosophila* leg imaginal disc^{1,2}, and it was proposed originally that they do this indirectly by defining an organizer in the centre: the presumptive tip of the leg, which would be the source of additional secreted signal(s)^{1,7}. However, more recent studies suggest that Wg and Dpp act more directly in proximodistal patterning and may function as morphogens to activate genes such as *Distal-less* (*Dll*) and *dachshund* (*dac*): the more Wg and Dpp a cell receives, the more distal it becomes³. At present, however, it is unclear how much of the proximodistal axis is actually specified directly by Wg and Dpp. This is important when considering appendage evolution because the legs of arthropod ancestors may largely correspond only to the distal region of *Drosophila* legs⁸, the tarsus, which constitutes most of the *Dll* domain (Fig. 1a, e). This was revealed by removing *Hox* gene influence in legs and antennae, the supposition being that this would produce a 'ground state' appendage that would reflect the nature of the leg in an arthropod ancestor before leg patterning was modified by *Hox* gene activity⁸. These ground-state appendages consist of a proximal region with an irregular morphology unlike that of an appendage, and a distal leg-like region consisting of an almost normally patterned tarsus.

Initially, we tested whether Wg and Dpp pattern directly the proximodistal axis of the tarsus by determining their role in activation of the *aristaless* (*al*) gene in the centre of the disc (Fig. 1a, b)¹. *al* encodes for a homeodomain protein required for development of structures found at the tip of the leg, including the claws¹³. Previous studies indicated that *al* expression is activated by Wg and Dpp¹ and this was confirmed with loss of function studies: *al* expression is absent from the centre of *wg* and *dpp* mutant discs (Fig. 1f, g). However, this does not rule out *al* being activated by a secondary signal, which in turn is activated by Wg and Dpp. To test this, *al* expression was monitored in discs containing clones of cells mutant for genes required for transduction of Wg or Dpp signals, including *arrow* (*arr*), which encodes for a Wg co-receptor⁹, and *thickveins* (*tkv*), which encodes for a Dpp receptor^{10,11} (clone founder cells were generated before the onset of *al* expression). Central *al* expression was absent in discs consisting largely of *arr* mutant clones (Fig. 1h), but, as in *wg*^{ts} discs, such large clones would remove any putative secondary signal, and, in fact, further analysis revealed that *al* was still expressed in *arr* mutant cells located outside of the very centre (Fig. 1i). Similarly, *al* could still be detected in *tkv* mutant cells (Fig. 1j). Thus, Wg and Dpp signalling are required, but not directly, to induce *al*, suggesting that it is activated by a secondary signal, which in turn is activated by Wg and Dpp.

The EGF-receptor (EGFR) signalling pathway was tested as a potential activator of *al* because the claws are lost in *Egfr* hypomorphic mutants¹² (the claws are also lost in *al* mutants¹³). Use of a temperature-sensitive (*ts*) allele¹⁴ showed that *Egfr* is required for development of almost all of the tarsus (Fig. 2b). After a 24-h shift to the restrictive temperature during the first half of the third instar, *Egfr*^{ts} adults have leg truncations with the size of the truncation increasing with temperature. At lower temperatures only the most distal structures, the claws, are lost, but at high temperatures tarsal segments II–V are absent. Intermediate truncations result at temperatures between these extremes (Fig. 2b). These shifts affect only patterning of the tarsus, apart from at the highest temperature, 33 °C, when development of more proximal regions is occasionally disrupted (Supplementary Information Fig. 2); this may reflect a low level EGFR signalling requirement for proliferation as has been shown in the eye disc¹⁵. The temperature-dependent truncations of the tarsus suggest that development of the most distal region

requires the highest level of EGFR activity, whereas more proximal regions require progressively less; that is, the tarsus may be patterned by a distal-to-proximal gradient of EGFR activity.

Similar results were obtained analysing marker gene expression in *Egfr*^{ts} discs, including *al*, *Bar* (*B*, expressed in segments IV and V; Fig. 1b)¹⁶ and *rotund* (*rn*, expressed in segments II–IV; Fig. 1c)¹⁷. Loss of EGFR activity results in loss of *al*, *B* and *rn* expression, but *al* is lost at a lower temperature than *B*, which in turn is lost at a lower temperature than *rn* (Fig. 2d–g), indicating that the more distal the marker, the higher the EGFR activity level required for expression. Clonal analysis with *Egfr*^{ts} showed that this response to EGFR is cell autonomous, and that again, *al* required higher EGFR activity than *B* (Fig. 2j, k). It was not possible to do similar tests for *rn* because at temperatures above 31 °C *Egfr*^{ts} clones do not survive in distal regions (not shown), raising the possibility that *rn* expression may be lost in *Egfr*^{ts} discs (Fig. 2g) simply because of reduced growth or cell survival in this region. However, expression of other genes, including *wg* and *dpp*, appears normal in *Egfr*^{ts} discs (Fig. 2h, i). This is also true for the Wg and Dpp target *Dll* (Fig. 2g), clearly demonstrating that Wg, Dpp and *Dll* are not sufficient to distalize the leg.

Ectopic activation of EGFR signalling results in autonomous, ectopic expression of *al* and *B* in mid-third-instar discs (Fig. 2m–p). Curiously, not all regions of the disc respond identically, with ventral regions being the most responsive and lateral regions the

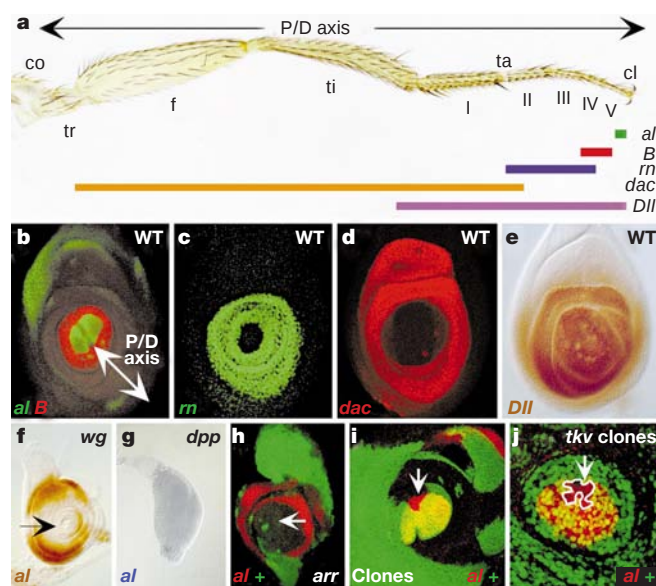


Figure 1 Proximodistal axis of the wild-type (WT) leg and regulation of *al* expression by Wg and Dpp signalling. **a**, The proximodistal (P/D) axis of the adult leg is divided into segments: co, coxa; tr, trochanter; f, femur; ti, tibia; ta, tarsus, which is divided into five segments (I–V) and terminates in a claw organ (cl). The expression domains of *al*, *B*, *rn*, *dac* and *Dll* in late third instars are indicated. **b–j**, Late third-instar leg discs. **b**, *al* is expressed in the centre, the presumptive tip, and *B* is expressed in a domain surrounding it (it is absent from the centre). *al* is also expressed in a more proximal domain. The proximodistal axis of the leg corresponds to the radius of the disc. **c**, *rn* is expressed in a wider domain than *B* and is absent from the centre. **d**, *dac* expression extends further proximally than *rn* or *Dll* and is absent from a wider region in the centre of the disc than *rn* or *B*. **e**, *Dll* is expressed throughout the distal region. **f**, *wg*^{ts} leg disc from a larva raised at the restrictive temperature during the third instar; *al* expression is absent from the centre (arrow), but is expressed proximally. **g**, No *al* expression is detected in *dpp* mutant discs. **h**, Disc containing *arr* mutant clones marked by the loss of a *B*-gal marker (green). Clones fill all of the central region and *al* expression is absent from there (arrow). **i**, *al* is still expressed (arrow) in *arr* mutant cells when these do not fill all of the centre. **j**, *al* is still expressed in *tkv* mutant cells (arrow, the clone is outlined). (Single channel images of Figs 1 and 2 are available in Supplementary Information).

least (Fig. 2v). The reason for this is unclear but it indicates that factors in addition to EGFR may be regulating expression of tarsal genes such as *al*, at least outside of their normal domains. Only the presumptive tarsus is responsive to ectopic EGFR activity; this is most evident in adult appendages where no defects in patterning can be observed outside of here (Fig. 2c). Other regions may be refractive to ectopic EGFR activity because expression of tarsal genes requires Dll^{13,16}, which is expressed only in distal regions under Wg and Dpp control^{2,3}.

In addition to activating genes, EGFR signalling is required to repress genes in distal regions, and again different genes appear to be differentially sensitive, with some, such as *B* and *rn*, possibly being both activated and repressed above different thresholds. *B*, *rn* and *dac* are repressed in the centre of wild-type discs, with *dac* being repressed over a wider region than *B* and *rn* (Fig. 1b, d). Lowering EGFR activity in *Egfr^{ts}* discs to a level sufficient only for loss of *al*, results in expression of *B* and *rn* in the centre, but not *dac* (Fig. 2d, e). Raising the temperature still further results in extension of the *dac* domain to fill the centre (Fig. 2f). Clonal analysis shows that

Egfr acts autonomously to repress *dac* (Fig. 2l). Ectopic EGFR activity can also repress *B*, *dac* and *rn* but again predominantly in ventral regions (*B* is repressed mainly at later stages; Fig. 2q–u). Previous studies have shown that repression of *dac* in distal regions requires high levels of Wg and Dpp signalling³, so all three pathways appear to be required to achieve this.

A distal-to-proximal gradient of EGFR activity predicts a source of ligand(s) at the presumptive tip. Potential ligands are the TGF- α family members Spitz and Keren, and the neuregulin, Vn; the former require activation by the membrane protein Rhomboid (Rho), or the homologue Roughoid (Ru)^{18–20}. Both *vn* and *rho* are expressed in the centre of the leg disc in early third instars (Fig. 3a, b). Genetic studies show that they are redundant so that loss of either gene alone has no effect on tarsus development (Supplementary Information Fig. 3), but loss of both together along with *ru*, which shows partial redundancy with *rho* (Supplementary Information Fig. 3) even though no expression can be detected (Fig. 3c), has marked effects on leg patterning and growth. Large *ru rho vn* triple mutant clones can result in truncations of the tarsus (Fig. 3e),

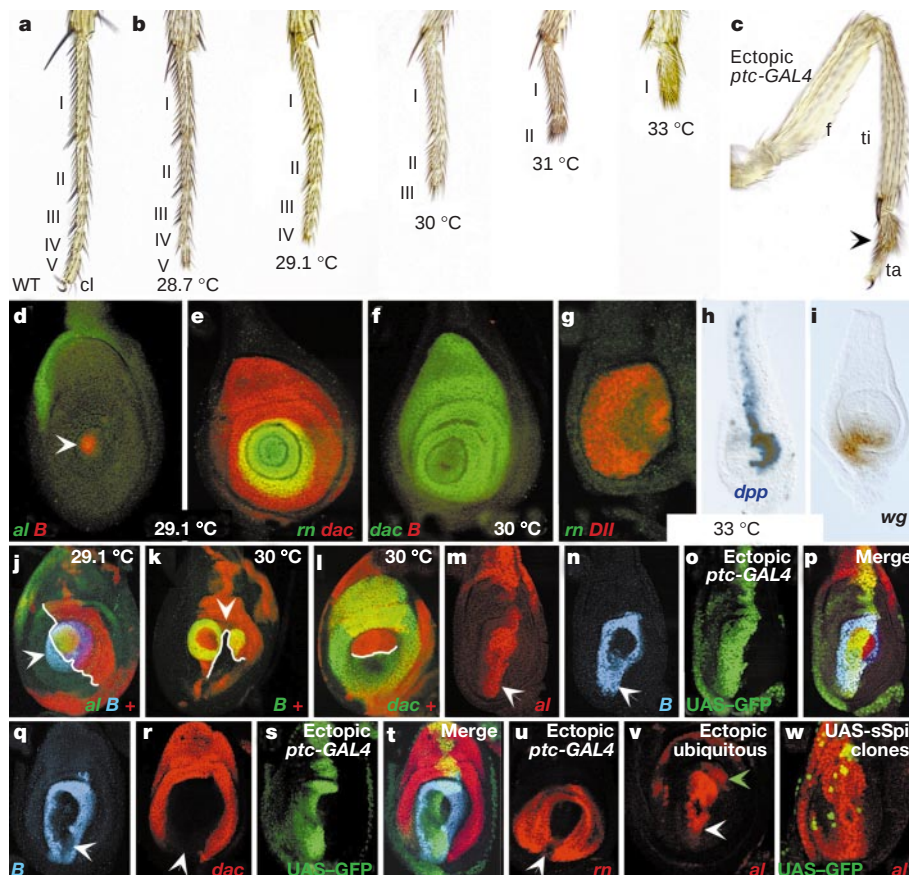


Figure 2 Regulation of tarsal development by EGFR signalling. **a**, Wild-type (WT) adult tarsus. **b**, Adult tarsi from *Egfr^{ts}* mutants raised at restrictive temperatures for the first half of the third instar. These have truncations, the severity increasing with temperature. **c**, Ectopic expression of a constitutively active form of EGFR using *ptc-GAL4* results in aberrant patterning of the tarsus (arrow), whereas the rest of the leg is normal. **d–i**, Late third-instar leg discs from *Egfr^{ts}* mutants raised at restrictive temperatures throughout the third instar. Patterns of gene expression are indicated. **d–e**, 29.1°C. **d**, Central *al* expression (arrow) is lost, while *B* is now expressed in the centre. **e**, *m* is also expressed in the centre, but *dac* is still absent. **f**, At 30°C, *B* expression is lost and *dac* extends into the centre. **g–i**, 33°C. **g**, *m* expression is lost, but *Dll* is still expressed strongly in the centre. **h, i**, *dpp* (**h**) and *wg* (**i**) expression appear normal. **j–l**, Discs containing *Egfr^{ts}* clones marked by the loss of a β -gal marker (red); the edge of mutant tissue is partially outlined. **j**, At 29.1°C, *al* expression is lost autonomously in *Egfr^{ts}* mutant cells, but *B* is still

expressed. **k**, At 30°C, *B* is lost autonomously. **l**, At 30°C, *dac* is autonomously misexpressed in *Egfr^{ts}* mutant cells, although this is very weak in some places. **m–v**, Ectopic activation of EGFR signalling as in **c** (monitored with UAS–GFP; **o, s, w**). **m–p**, Mid-third-instar disc showing autonomous ectopic *al* (red, **m, p**) and *B* (blue, **n, p**) expression, predominantly in ventral regions (arrows). **q–t**, Late third instar disc showing repression of *B* (blue, **q, t**) and *dac* (red, **r, t**) in ventral regions (arrows; *dac* is only repressed in the presumptive tarsus, more proximal expression (not shown) is normal). **u**, *m* is also repressed in ventral regions (arrow). **v**, Ubiquitous activation of EGFR signalling with *tubulin-GAL4* shows ectopic activation of *al* predominantly ventrally (white arrow), some dorsally (green arrow), but very little elsewhere. **w**, Clones of cells (UAS–GFP) misexpressing a secreted form of Spitz activate *al* expression in surrounding cells (again predominantly dorsally and ventrally; these clones do not grow well in distal regions and some may be lost by this stage).

although these are never as extreme as in *Egfr^{ts}* mutants, possibly because of the difficulty of removing all ligand-expressing cells at the centre of early leg discs using this technique, or because the *ru* mutant used is not null. Wild-type tissue located at the tip of adult legs always correlates with rescue of tarsal development (Fig. 3f). In addition, misexpression of a secreted form of Spitz results in non-autonomous activation of *al* (Fig. 2w). Verification of high levels of EGFR signalling in the distal leg is revealed by expression of *sprouty* in this location (Fig. 3d); this is upregulated in many tissues by EGFR signalling^{21,22}.

Although a quantitative response to EGFR activity is important, an additional factor that can influence patterning of the tarsus is timing, revealed by shorter temperature shifts with the *Egfr^{ts}* mutant. A 12-h shift to 33 °C in the first quarter of the third instar also results in leg truncations (Fig. 4a), although these are less severe than those produced by the 24-h shifts described above. However, a similar 12-h shift later during the second quarter of the third instar results in legs with normal distal pattern elements (claw, tarsal segment V), but with fusions or deletions of intermediate segments II–IV (Fig. 4b), suggesting that distal regions are specified before more proximal ones. This phenomenon mirrors the temporal pattern of gene expression in these regions in wild-type discs: the distal markers *al* and *B* are expressed in very early third instars whereas the intermediate tarsal marker, *rn*, cannot be detected until several hours later (Fig. 4c, d)¹⁷. The early 12-h shift can prevent all *al* and *B* expression (Fig. 4e), correlating with the pattern defect.

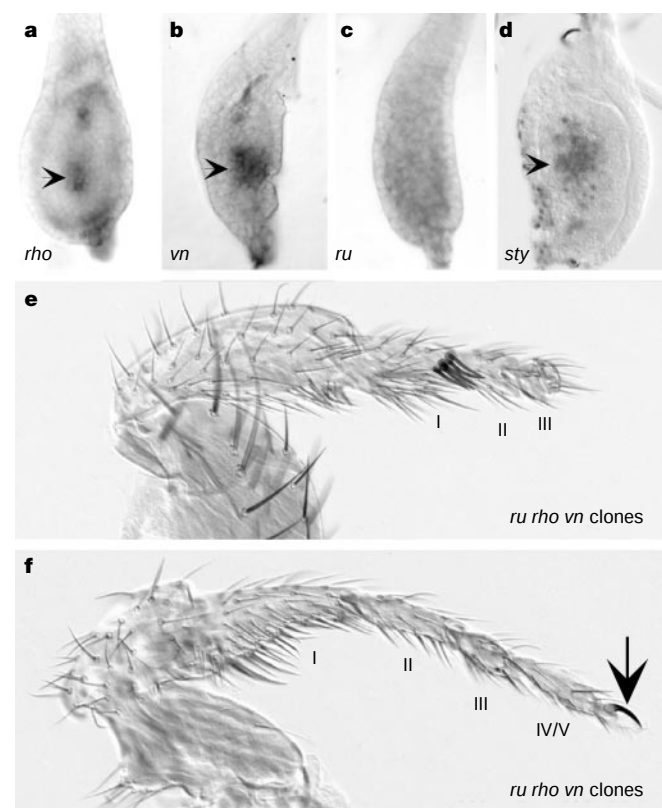


Figure 3 Expression and activity of EGFR ligands in the tarsus. **a–c**, *In situ* of early third instar discs from wild-type larvae. Both *rho* (**a**) and *vn* (**b**) are expressed in the centre, but no *ru* expression can be detected (**c**). **d**, *sty* is also expressed in the centre of discs at this stage. **e, f**, Adult legs comprised almost exclusively of *ru rho vn* triple mutant clones. **e**, The tarsus is truncated with only three segments. Development of more proximal tissue is also disrupted, possibly a reflection of a low level EGFR requirement for proliferation (*ru* and *vn* are expressed more proximally at later stages). **f**, In this leg there is a small region of wild-type tissue at the tip (arrow; the claw is the only detectable *y⁺* marker in this leg), which has rescued development of the tarsus to almost five segments.

However, the later shift does not result in loss of *rn* expression (although the domain is smaller than in wild-type discs; Fig. 4i). It is possible that the patterning defect resulting from the later shift may be caused by the loss of other, as yet unidentified, EGFR targets at this time. Alternatively, growth and patterning of this region requires segmentation of the tarsus, which is controlled by Notch signalling²³, so it is possible that EGFR signalling may be involved in regulating this process in the tarsus. In addition, these shifts are associated with cell death (Fig. 4k), which may also contribute to the defective patterning.

The present results are consistent with the proposal that the tip of the leg is the source of EGF-related ligands and acts as an organizer to regulate proximodistal patterning of the tarsus, the evolutionarily 'ancient' portion of the *Drosophila* leg⁸. This does not rule out a role for Wg and Dpp, but these proteins appear to be required only indirectly for regionalization of the tarsus, first to activate *Dll* during the first/second instar^{2,3} and then probably to activate *rho* and *vn* in the centre of the disc. The resulting EGFR gradient is then used to establish the patterns of tarsal gene expression during the third instar. This can be compared to proximodistal patterning in vertebrates where the apical ectodermal ridge at the tip of the limb bud is the source of FGFs that are required for outgrowth and formation of the proximodistal axis⁴—both EGF and FGF receptors are RTKs. This similarity could indicate an evolutionary link between outgrowths from the body wall in animals, a possibility already revealed by the expression of *Dll* homologues in most of these outgrowths⁵. This is not to imply that vertebrate and insect limbs evolved directly from a putative body-wall outgrowth in their last common ancestor, but that this putative outgrowth probably used a developmental genetic pathway⁶ involving *Dll* and distally expressed RTK ligands, which has subsequently been used in new ways in these two phyla and possibly elsewhere in the animal kingdom. □

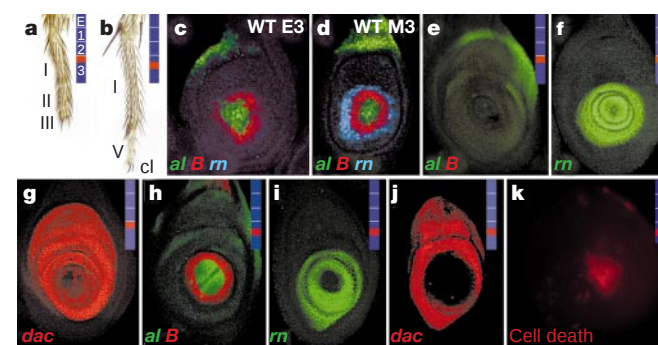


Figure 4 Temporal requirement for EGFR signalling. **a, b**, Adult tarsi from *Egfr^{ts}* mutants raised at 33 °C for 12 h for the first quarter (**a**) and the second quarter (**b**) of the third instar. The upper right panel indicates time (E, embryo; 1–3, first to third larval instars) spent at 18 °C (blue) and at 33 °C (red). The earlier shift results in truncations of the tarsus (**a**), whereas the later shift results in normal distal-most structures (claws and segment V), but the intermediate segments (II–IV) are lost (**b**). **c, d**, *al*, *B* and *rn* expression in early third-instar (**c**) or mid-third-instar (**d**) leg discs. *al* and *B*, but not *rn* are expressed at early stages. By mid-third instar, *rn* can be detected and is mostly expressed in a ring surrounding *B* (at later stages there is considerable overlap). **e–g**, Late third instar discs from *Egfr^{ts}* larvae raised at 33 °C for 12 h in the first quarter of the third instar. **e**, *al* and *B* expression are lost in the centre. **f, g**, *m* (**f**) and *dac* (**g**) expression extends into the centre. **h–k**, Same as for **e–g**, but the shift was during the second quarter of the third instar. **h**, *al* and *B* are expressed as in wild type. **i**, *rn* expression is also expressed in its normal domain, although this appears reduced compared with wild type. **j**, *dac* expression appears fairly normal. **k**, This disc was fixed 24 h after returning the larvae to the permissive temperature and stained with acridine orange to detect dying cells. These could be found in the distal leg, but they had delaminated from the epidermal sheet.

Methods

Egfr, *wg* and *dpp* loss and gain of function

Late third-instar discs were dissected from the *wg* temperature-sensitive mutant, *wg*¹⁻¹², which was shifted to the restrictive temperature (29 °C) for all of the third instar, and from the *dpp* disc-specific mutant, *dpp*^{d12/d14}. *Egfr*^{ts} refers to the genotype *Egfr*^{tsla}/*Egfr*^{ts4} (ref. 14). For the initial adult analysis, larvae with this genotype were raised at 18 °C apart from a 24-h shift to the restrictive temperature commencing during the early third instar. For the initial disc analysis, larvae were maintained at the restrictive temperature throughout the third instar and discs were dissected and fixed at the end of this period. For the later timing experiments, 12-h shifts were made at the beginning of the third instar or 24 h later (third instar duration of 4 days at 18 °C) and discs were fixed from late third instars or larvae were allowed to develop to adults.

Ectopic activation of the EGFR pathway was achieved using the UAS/Gal4 system with flies carrying the transgene *UAS-Egfr*^{tsla}, which codes for a constitutively active form²⁴, and *ptc*-GAL, which is expressed along the anteroposterior compartment boundary (at fairly uniform levels in dorsal and ventral halves); this expression was monitored with UAS-green fluorescent protein (GFP). *UAS-Egfr*^{tsla} was also expressed ubiquitously using the flip-out technique²⁷ with an *hs-flp*; tubulin > CD2 > Gal4 line: larvae were given a 37 °C heat shock for 1 h during the second instar, to induce ubiquitous expression. Misexpression of a secreted form of Spitz, *UAS-sSpi*, was achieved in a similar manner, but larvae were given a much shorter heat shock of 32.5 °C for 30 min to generate small clones.

Gene expression

Immunostaining was performed using standard techniques with antibodies against the following proteins: Al (rat)¹, Dll (mouse)², B (rabbit)¹⁶, Wg (mouse)²⁵, Dac (mouse)²⁶, β -galactosidase (rabbit; Cappel), β -galactosidase (mouse; Promega). *dpp* and *sty* expression was detected with enhancer traps (staining with an antibody against activated mitogen-activated protein kinase is not reliable in this tissue). *vn*, *rho* and *ru* expression was detected by *in situ* hybridization using standard techniques (unfortunately, antibodies or reliable lacZ lines are not available to test currently whether *vn* and *rho* expression requires autonomous activation of Wg and Dpp signalling pathways).

Clonal analysis

Loss of function clones were generated by flip-mediated mitotic recombination, in conjunction with the minute technique, using the following stocks: *FRT42 arr²/Cyo*, *FRT39E tkv⁷/Cyo*, *FRT42 Egfr^{tsla}/Cyo*, *y; rho^{ts45} FRT80B/TM6B*, *y; ru¹ rho^{TM43} FRT80B/TM6B*, *y; ru¹ rho^{TM43} vn^{LC} FRT2A/TM6B*, which were crossed to one of the following marker stocks: *hs-flp; FRT42 arm-lacZ M(2)60E/Cyo*, *hs-flp; hsGFP M(2)201 FRT39E/Cyo*, *y hs-flp; y⁺ M(3)i³⁵ FRT80B/TM2*, *y hs-flp; y⁺ M(3)i³⁵ FRT2A/TM2*, *hs-flp; hsGFP M(3)i³⁵ FRT2A/TM6B*. The *arr*, *tkv*, *vn* and *rho* alleles are genetically or molecularly null, *ru¹* is a hypomorph. Clones were induced during the second instar and discs were dissected and fixed from late third instars. Disc clones were marked by the loss of GFP or lacZ transgenes; adult clones were marked by the loss of *y⁺*. More details on all these alleles can be found at FlyBase (<http://flybase.bio.indiana.edu>).

Received 8 May; accepted 11 July 2002; doi:10.1038/nature00971.

Published online 24 July 2002.

- Campbell, G., Weaver, T. & Tomlinson, A. Axis specification in the developing *Drosophila* appendage: the role of wingless, decapentaplegic, and the homeobox gene *aristaless*. *Cell* **74**, 1113–1123 (1993).
- Diaz-Benjumea, F. J., Cohen, B. & Cohen, S. M. Cell interaction between compartments establishes the proximal-distal axis of *Drosophila* legs. *Nature* **372**, 175–179 (1994).
- Lecuit, T. & Cohen, S. M. Proximal-distal axis formation in the *Drosophila* leg. *Nature* **288**, 139–145 (1997).
- Martin, G. R. The roles of FGFs in the early development of vertebrate limbs. *Genes Dev.* **12**, 1571–1586 (1998).
- Panganiban, G. et al. The origin and evolution of animal appendages. *Proc. Natl Acad. Sci. USA* **94**, 5162–5166 (1997).
- Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
- Campbell, G. & Tomlinson, A. Initiation of the proximodistal axis in insect legs. *Development* **121**, 619–628 (1995).
- Casares, F. & Mann, R. S. The ground state of the ventral appendage in *Drosophila*. *Science* **293**, 1477–1480 (2001).
- Wehrli, M. et al. *arrow* encodes an LDL-receptor-related protein essential for Wingless signalling. *Nature* **407**, 527–530 (2000).
- Nellen, D., Affolter, M. & Basler, K. Receptor serine/threonine kinases implicated in the control of *Drosophila* body pattern by decapentaplegic. *Cell* **78**, 225–237 (1994).
- Penton, A. et al. Identification of two bone morphogenetic protein type I receptors in *Drosophila* and evidence that Brk25D is a decapentaplegic receptor. *Cell* **78**, 239–250 (1994).
- Clifford, R. J. & Schupbach, T. Coordinately and differentially mutable activities of torpedo, the *Drosophila melanogaster* homolog of the vertebrate EGF receptor gene. *Genetics* **123**, 771–787 (1989).
- Campbell, G. & Tomlinson, A. The roles of the homeobox genes *aristaless* and *Distal-less* in patterning the legs and wings of *Drosophila*. *Development* **125**, 4483–4493 (1998).
- Kumar, J. P. et al. Dissecting the roles of the *Drosophila* EGF receptor in eye development and MAP kinase activation. *Development* **125**, 3875–3885 (1998).
- Halfar, K., Rommel, C., Stocker, H. & Hafen, E. Ras controls growth, survival and differentiation in the *Drosophila* eye by different thresholds of MAP kinase activity. *Development* **128**, 1687–1696 (2001).
- Kojima, T., Sato, M. & Saigo, K. Formation and specification of distal leg segments in *Drosophila* by dual Bar homeobox genes, BarH1 and BarH2. *Development* **127**, 769–778 (2000).
- Couso, J. P. & Bishop, S. A. Proximo-distal development in the legs of *Drosophila*. *Int. J. Dev. Biol.* **42**, 345–352 (1998).
- Schweitzer, R. & Shilo, B. Z. A thousand and one roles for the *Drosophila* EGF receptor. *Trends Genet.* **13**, 191–196 (1997).

- Bang, A. G. & Kintner, C. Rhomboid and Star facilitate presentation and processing of the *Drosophila* TGF- α homolog Spitz. *Genes Dev.* **14**, 177–186 (2000).
- Wasserman, J. D., Urban, S. & Freeman, M. A family of rhomboid-like genes: *Drosophila* rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling. *Genes Dev.* **14**, 1651–1663 (2000).
- Kramer, S., Okabe, M., Hacohen, N., Krasnow, M. A. & Hiromi, Y. Sprouty: a common antagonist of FGF and EGF signaling pathways in *Drosophila*. *Development* **126**, 2515–2525 (1999).
- Casci, T., Vinos, J. & Freeman, M. Sprouty, an intracellular inhibitor of Ras signaling. *Cell* **96**, 655–665 (1999).
- Irvine, K. D. Fringe Notch, and making developmental boundaries. *Curr. Opin. Genet. Dev.* **9**, 434–441 (1999).
- Queenan, A. M., Ghabrial, A. & Schupbach, T. Ectopic activation of torpedo/Egfr, a *Drosophila* receptor tyrosine kinase, dorsalizes both the eggshell and the embryo. *Development* **124**, 3871–3880 (1997).
- Brook, W. J. & Cohen, S. M. Antagonistic interactions between wingless and decapentaplegic responsible for dorsal-ventral pattern in the *Drosophila* leg. *Science* **273**, 1373–1377 (1996).
- Mardon, G., Solomon, N. M. & Rubin, G. M. dachshund encodes a nuclear protein required for normal eye and leg development in *Drosophila*. *Development* **120**, 3473–3486 (1994).
- Struhl, G. & Basler, K. Organizing activity of wingless protein in *Drosophila*. *Cell* **72**, 527–540 (1993).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

I thank A. Tomlinson, M. Simcox, D. Chapman and R. Carthew for advice and comments on the manuscript. I also thank the following for stocks and reagents: K. Moses, J. Kumar, M. Werhli, A. Simcox, J. P. Couso, S. Kunes, T. Schubach, T. Volk, G. Struhl, H. M. Chung, M. Feeman, R. Cagan, T. Kojima, S. Cohen, K. Cook, K. Matthews, the Bloomington Stock Center, J. P. Vincent, and the Developmental Studies Hybridoma Bank.

Competing interests statement

The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (e-mail: camp@pitt.edu).

Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria

Louis Schofield*, Michael C. Hewitt†, Krystal Evans*, Mary-Anne Siomos* & Peter H. Seeberger†

* Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Victoria 3050, Australia

† Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

The malaria parasite *Plasmodium falciparum* infects 5–10% of the world's population and kills two million people annually¹. Fatalities are thought to result in part from pathological reactions initiated by a malarial toxin. Glycosylphosphatidylinositol (GPI) originating from the parasite has the properties predicted of a toxin^{2–6}; however, a requirement for toxins in general and GPI in particular in malarial pathogenesis and fatality remains unproven. As anti-toxic vaccines can be highly effective public health tools, we sought to determine whether anti-GPI vaccination could prevent pathology and fatalities in the *Plasmodium berghei*/rodent model of severe malaria. The *P. falciparum* GPI glycan of the sequence NH₂-CH₂-CH₂-PO₄-(Man α 1-2)6Man α 1-2Man α 1-6Man α 1-4GlcNH₂ α 1-6myo-inositol-1,2-cyclic-phosphate was chemically synthesized, conjugated to carriers, and used to immunize mice. Recipients were substantially protected against malarial acidosis, pulmonary oedema, cerebral syndrome and fatality. Anti-GPI antibodies neutralized pro-inflammatory activity by *P. falciparum* *in vitro*. Thus, we show that GPI is a significant pro-inflammatory endotoxin of parasitic origin, and that several disease parameters in malarious mice are toxin-dependent. GPI may contribute to pathogenesis and fatalities

in humans. Synthetic GPI is therefore a prototype carbohydrate anti-toxic vaccine against malaria.

Malarial GPI is a candidate toxin that is sufficient to induce cytokine and adhesin expression in macrophages and the vascular endothelium^{2,4–6}—both of which are associated with clinically severe malaria^{7,8}—and to induce lethality *in vivo*^{2–6}. GPIs of *Trypanosoma brucei*^{3,9} and *T. cruzi*¹⁰ have similar properties, suggesting that GPIs may act generally as pro-inflammatory agents in eukaryotic parasitism. However, it is not yet established whether GPIs function as toxins in the context of protozoal infections, nor whether intervention against GPIs reduces pathogenesis or fatalities in any disease condition. Indeed the toxic basis of malarial pathogenesis, first conjectured¹¹ by Camillo Golgi in 1886, remains unproven.

Plasmodium falciparum shows uniquely low levels of N- and O-linked glycosylation^{12,13}, and the highly conserved¹⁴ GPI constitutes over 95% of the post-translational carbohydrate modification of parasite proteins¹⁵. The biological activity of GPI against host tissues requires the contribution of both lipid and carbohydrate domains, and de-acylation of GPIs by enzymatic or chemical hydrolysis renders the carbohydrate moiety non-toxic^{2–6}. On the basis of the sequence of the non-toxic *P. falciparum* GPI glycan¹⁶, we chemically synthesized the structure NH₂-CH₂-CH₂-PO₄-(Man α 1-2)6Man α 1-2Man α 1-6Man α 1-4GlcNH₂ α 1-6myo-inositol-1,2-cyclic-phosphate (Fig. 1). We confirmed the structure by matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry and ³¹P-NMR (D₂O) (see Supplementary Information). To prepare an immunogen, the synthetic GPI glycan was treated with 2-iminothiolane to introduce a sulphhydryl at the ethanolamine, desalted, and conjugated to maleimide-activated ovalbumin (OVA), in a molar ratio of 3.2:1, or keyhole limpet haemocyanin (KLH), in a molar ratio of 191:1. This material was used to immunize mice.

The synthetic malarial GPI glycan was immunogenic in rodents. Antibodies from animals immunized with KLH-glycan gave positive immunoglobulin- γ (IgG) titres against OVA-glycan but not sham-conjugated OVA-cysteine. No reactivity to GPI glycan was detected in pre-immune sera or in animals receiving sham-conjugated KLH (not shown). Notably, anti-glycan IgG bound to native GPI, as judged by immunofluorescence against intact trophozoites and schizonts (Fig. 2a). Anti-GPI however failed to bind to uninfected erythrocytes, despite these cells expressing endogenous GPIs of host origin (Fig. 2a). In contrast to malarial GPI, mammalian GPIs show amino-sugar or phosphoethanolamine modifications to the core glycan¹⁷, and these epitopic differences may account for the lack of cross-reactivity. These results do not exclude the possibility of serological cross-reactions with other tissues. Unlike controls, in a western blot analysis anti-GPI glycan IgG detected multiple molecular species against *P. falciparum*-infected but not uninfected erythrocytes (Fig. 2b), consistent with the presence in mature schizonts of multiple GPI-modified proteins and their processing products. Thus protein-specific features do not greatly influence the binding of anti-glycan IgG to native GPI anchors.

Tumour-necrosis factor (TNF)- α production by macrophages is widely used as a biochemical marker of malarial endotoxin activity *in vitro*. Purified GPIs are sufficient for TNF production^{2–6,9,10}, but a predominant role for GPI in parasite pro-inflammatory activity *in vitro* remains unproven. We therefore sought to quantify the contribution of GPI to the total endotoxic activity of malaria. In contrast to control sera, antibodies from mice immunized with KLH-glycan specifically neutralized TNF- α output from macrophages induced by crude total extracts of *P. falciparum* (Fig. 2c). Thus GPI appears sufficient and necessary for the induction by malarial parasites of host pro-inflammatory responses *in vitro*. Naturally, other entities of host or parasite origin may also influence such responses.

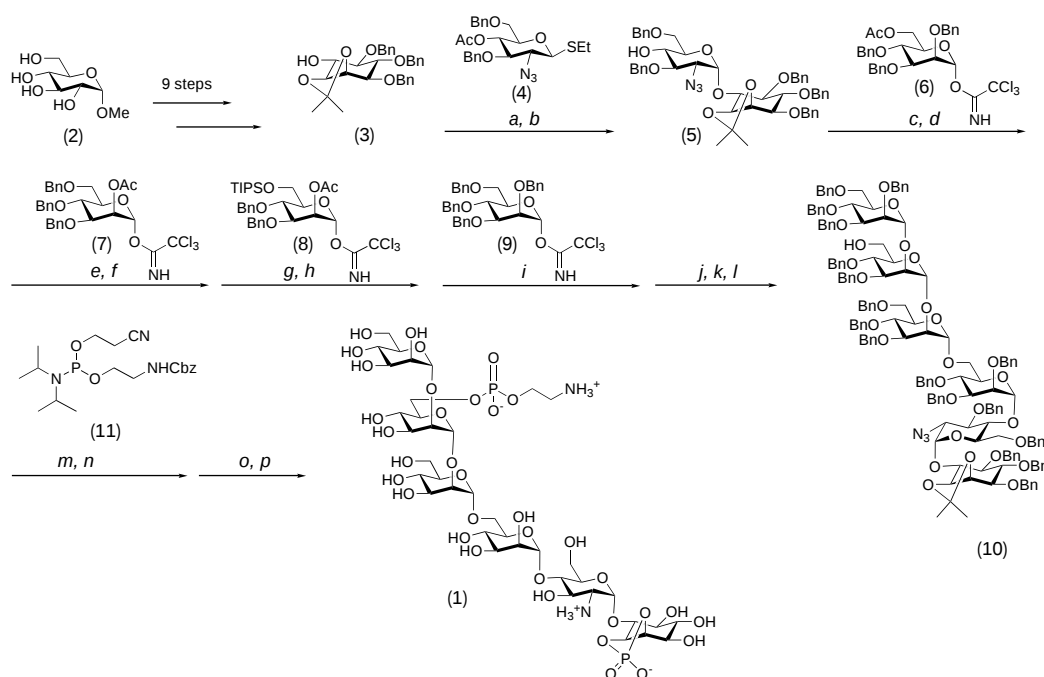


Figure 1 Synthesis of glycan (1). The reagents (a–p) were added at the indicated intermediates of glycan synthesis (1–10). a, (4) AgOTf, NIS, CH₂Cl₂/Et₂O (38%); b, NaOMe, CH₂Cl₂/MeOH (83%); c, (6) TMSOTf, CH₂Cl₂ (75%); d, NaOMe, CH₂Cl₂/MeOH (71%); e, (7) TMSOTf, CH₂Cl₂ (92%); f, NaOMe (69%); g, (8) TBSOTf, CH₂Cl₂ (98%); h, NaOMe (83%); i, (9) TMSOTf, CH₂Cl₂ (84%); j, (CH₂OH)₂, CSA, CH₃CN (81%); k, Cl₂P(O)OMe, pyridine (88%); l, TBAF, THF (61%); m, (11) tetrazole, CH₃CN; n, *t*-BuOOH, CH₃CN (84%, 2 steps); o, DBU, CH₂Cl₂; p, Na, NH₃, THF (75%, 2 steps). AgOTf, silver

trifluoromethanesulphonate; NIS, *N*-iodosuccinimide; CH₂Cl₂, dichloromethane; Et₂O, diethyl ether; NaOMe, sodium methoxide; MeOH, methanol; TMSOTf, trimethylsilyl trifluoromethane sulphonate; TBSOTf, *tert*-butyldimethylsilyl trifluoromethanesulphonate; CSA, camphorsulphonic acid; CH₃CN, acetonitrile; Cbz, carbobenzyloxy; TBAF, tetrabutylammonium fluoride; THF, tetrahydrofuran; DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; OBn, O-benzyl.

Humans that are affected by and dying of malaria may suffer systemic, single- or multi-organ involvement, including acute respiratory distress, coagulopathy, shock, acidosis, hypoglycaemia, renal failure, pulmonary oedema and neurological signs¹⁸. The murine *P. berghei* ANKA severe malaria model has salient features in common with several aspects of the human severe and cerebral malaria syndromes. It manifests a cytokine-dependent encephalopathy associated with upregulation of adhesins on the cerebral microvascular endothelium and attendant neurological complications^{19–22}. Pulmonary oedema, lactic acidosis, coagulopathies, shock and renal impairment are also observed²³. Unlike some, but not all, human cerebral cases of malaria, there is a macrophage infiltrate and compromised blood–brain barrier in the terminal or agonal stages of the murine syndrome. Nonetheless in the proximal or developmental stages the murine disease reflects more accurately the cytokine-dependent inflammatory cascade leading to cerebral and systemic involvement in humans, and thus seems the best available small animal model of clinically severe malaria^{24,25}. Validation of GPI as a toxin and a target in this model might therefore allow the development of anti-toxic vaccines and immunotherapeutics that are able to prevent pathogenesis and fatalities in humans.

To this end, C57BL/6J mice primed and boosted twice with 6.5 µg KLH-glycan (0.18 µg glycan) or KLH-cysteine in Freund's adjuvant were challenged with *P. berghei* ANKA. All sham-immunized and naive control mice died with the cerebral syndrome, showing severe neurological signs including loss of reflex, ataxia, and hemiplegia, with hypothermia and occasional haematouria (Fig. 3a). These fatalities were evident early during infection (day 5–8), with relatively low levels of parasitaemia. There were no differences between naive and sham-immunized mice, indicating that exposure to KLH in Freund's adjuvant does not influence the rate of disease. In contrast, mice immunized with chemically synthetic *P. falciparum* GPI glycan coupled to KLH were significantly protected against severe malaria, with clearly reduced death rates (75% survival, $P < 0.02$, Fig. 3a). In four separate additional experiments, results in the range of 58.3–75% survival to day 12 in vaccine recipients ($n = 50$ total) compared with 0–8.7% survival in sham-immunized controls ($n = 85$) were obtained. Parasitaemia levels were not significantly different between test and control groups, demonstrating that prevention of fatality by anti-GPI vaccination does not operate through effects on parasite replication (Fig. 3b). The diagnoses of cerebral malaria, or absence of this condition, were confirmed by histological examination of brains taken 6 days after infection. Sham-immunized mice showed typical pathology including vascular occlusion with both parasitized red blood cells and host leukocytes (Fig. 3c). Immunized animals in contrast showed absent or reduced vascular occlusion despite similar parasite burdens (Fig. 3c).

Severe malaria in both humans¹⁸ and rodents²³ may be associated with additional organ-specific and systemic derangements, including pulmonary oedema and acidosis. Acidosis may be a prime pathophysiological process and is the strongest single prognostic indicator of fatality. The biochemical aetiology of acidosis is unclear, and the relationship of human malarial acidosis to that in the rodent model also remains to be elucidated. Nonetheless, we sought to determine whether anti-GPI vaccination protects against these additional non-cerebral disease syndromes in mice. Both sham-immunized and naive individuals developed pulmonary oedema by day 6 after infection, as measured by lung dry/wet weight ratios, and this was markedly reduced in vaccine recipients (Fig. 3d). Similarly, whereas sham-immunized and naive mice developed significant acidosis as shown by reduced blood pH at day 6, blood pH was maintained at physiological levels in mice that had received the vaccine (Fig. 3e). As parasite burdens were similar in both test and control groups, production of lactic acid by parasite biomass—and any haemolytic anaemia due to parasitaemia at this stage—are not

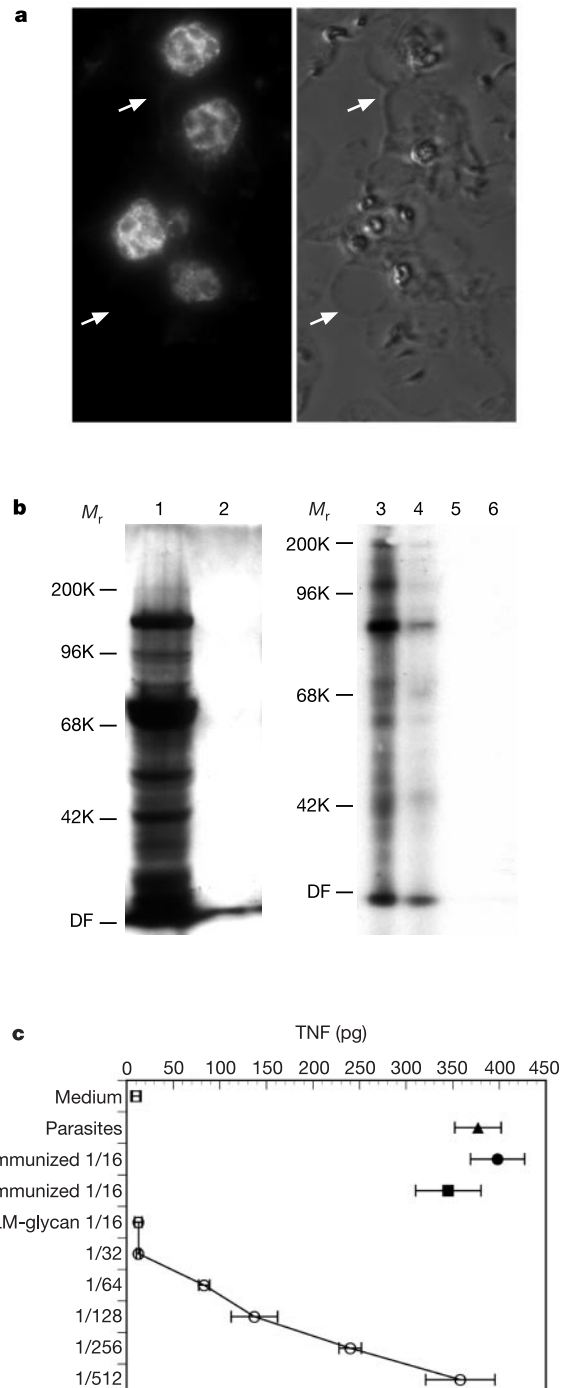


Figure 2 Antibodies raised against synthetic GPI glycan recognize native GPI and neutralize toxin activity *in vitro*. **a**, Reactivity of anti-glycan IgG antibodies with *P. falciparum* trophozoites and schizonts, and lack of reactivity to uninfected erythrocytes detected by immunofluorescence assay (left panel). The right panel shows the same field under white-light illumination. Arrows indicate adjacent uninfected erythrocytes. **b**, Western blot of anti-glycan IgG antibodies (1/200) against parasite-infected (lane 1) and uninfected erythrocytes (lane 2) run in 20-cm slab gel (left panel). The right panel shows a mini-blot comparison of reactivity against parasites by two sera from KLH-glycan-immunized mice (lanes 3, 4), pre-immune serum from lane 3 donor (lane 5), and serum from a mouse immunized with sham KLH (lane 6). All sera were used at 1/400 dilution. The detection antibody was peroxidase-conjugated goat anti-mouse IgG (γ -chain specific). DF, dye front; M_r , relative molecular mass. **c**, Levels ($\pm$ s.e.m.) of TNF- α in culture supernatants of RAW264.7 cells exposed in triplicate to medium alone (open square), parasites alone (triangle), or parasites in the presence of various dilutions of sera from pre-immune (filled circle), sham-immunized (filled square) or glycan-immunized mice (open circles).

major contributors to acidosis in this model. Clearly, immunizing against GPI prevents the development of pulmonary oedema and acidosis as well as cerebral malaria in *P. berghei* infection.

The aetiology of malarial anaemia in humans is complex and poorly understood. A principal contributory factor is thought to be the failure of stem cells in the bone marrow to repopulate the peripheral erythrocyte compartment, a process known as dyserythropoiesis or erythropoietic suppression. Although *P. berghei* is the best available model for certain aspects of lethal pathogenesis, it is not considered to model adequately these aspects of human malarial anaemia. Indeed, infection models of this condition are not yet fully developed. Unlike malarial infection in humans, *P. berghei* invariably proceeds to a late-end-stage infection characterized by overwhelming parasitaemia associated with profound haemolytic anaemia. Anti-GPI vaccination did not prevent this process, as all immunized animals eventually succumbed to massive parasitaemias by day 15 (mean $64.5\% \pm 12.1$), associated with a 75% reduction in erythrocyte density (data not shown). Although anti-GPI vaccination did not prevent hyperparasitaemia with attendant haemolytic anaemia, the relevance of these observations to human parasite burdens and dyserythropoietic anaemia remains unclear.

This study was designed to test the hypothesis that GPI is causally involved in rodent malarial pathogenesis, including metabolic derangement, and to determine whether vaccination against this target affords clinical protection in the best small animal model available. Mice were primed and boosted with 176 ng glycan per dose, which may be a suboptimal quantity. Systematic optimization with respect to formulation, carrier/hapten ratios, adjuvants, and

dosage or timing of the immunization regimen is beyond the scope of this study. Therefore it is possible that the degree of protection against disease observed here may improve further depending on these variables. Similarly, anti-GPI vaccination may conceivably be beneficial in other malarial disease syndromes not sufficiently modelled by acute *P. berghei* ANKA infection, for example, dyserythropoietic anaemia.

After initial susceptibility to severe disease, children in holoendemic regions are thought to develop acquired clinical immunity that protects against life-threatening pathology despite persistent high levels of parasitaemia^{26–28}. The validity of this proposition, and whether GPI is a target of clinical immunity, remain to be determined. GPI may be non-self in humans, and antibodies to GPI lipid domains may be associated with protection against disease²⁹. The chemical synthesis of GPI fragments reported here should aid in testing these hypotheses and in epitope mapping of human anti-GPI antibodies. In contrast to acquired clinical immunity, anti-parasite immunity takes many more years to develop²⁸ and is easily lost, reflecting the problems of antigenic diversity, antigenic variation, redundancy in invasion pathways, immune evasion strategies and genetic restriction in the immune response to parasite antigens. Current approaches to anti-malarial vaccines seek nonetheless to induce anti-parasite immunity through parasitocidal mechanisms targeted to parasite protein antigens. The public health potential of alternative anti-disease vaccine strategies is demonstrated by the highly effective tetanus and diphtheria toxoid vaccines that protect against the most injurious consequences of infection by targeting bacterial toxins³⁰. The findings of this study suggest that GPI is a highly conserved endotoxin of malarial parasite origin. A non-toxic

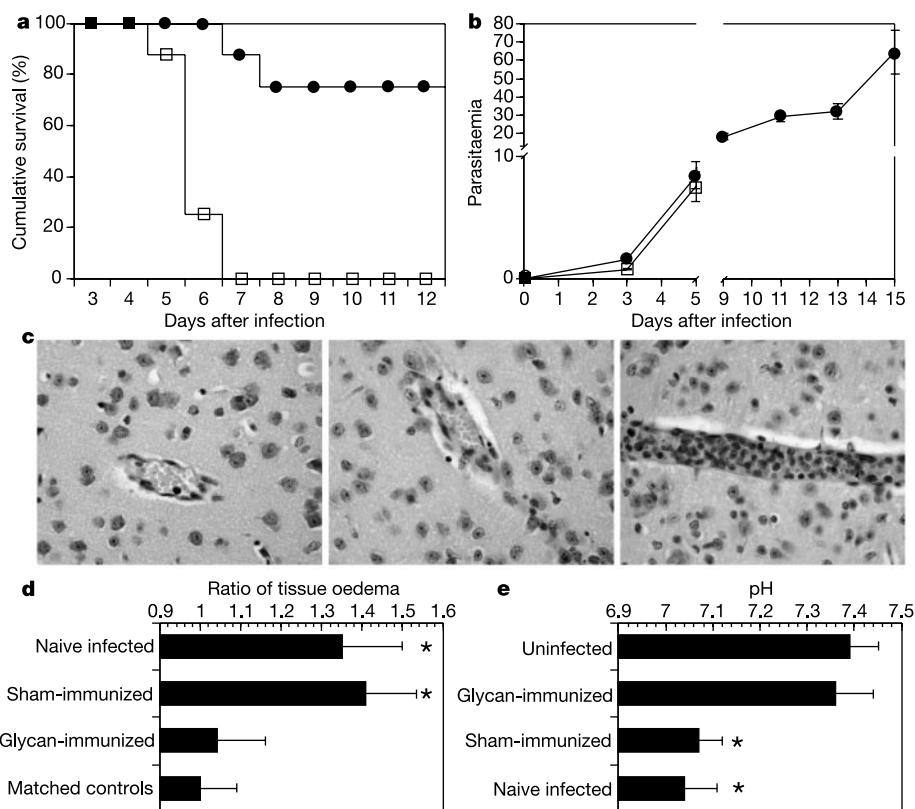


Figure 3 Immunization against the synthetic GPI glycan substantially protects against murine cerebral malaria, pulmonary oedema and acidosis. **a**, **b**, Kaplan–Meier survival plots (**a**) and parasitaemia levels (number of parasites per 100 red blood cells, **b**) of KLH-glycan-immunized (filled circles) and sham-immunized (open squares) mice challenged with *P. berghei* ANKA. **c**, Haematoxylin and eosin-stained sections of brain tissue showing blood vessels from KLH-glycan-immunized (left and centre panels) and sham-immunized (right panel) mice killed on day 6 after infection. **d**, As an index of pulmonary oedema, the

ratio of wet weight to dry weight of lungs from KLH-glycan-immunized ($n = 5$), sham-immunized ($n = 7$) and naive mice ($n = 8$) at day 6 after infection are expressed as a proportion of the lung wet:dry weight ratio of age/sex-matched uninfected ($n = 5$) controls. **e**, pH ($\pm$ s.e.m.) of serum drawn at day 6 from uninfected mice ($n = 4$) and from naive ($n = 6$), immunized ($n = 5$) and sham-immunized ($n = 6$) donor mice infected with *P. berghei* ANKA. Asterisk, $P < 0.05$.

GPI oligosaccharide coupled to carrier protein is immunogenic and provides significant protection against malarial pathogenesis and fatalities in a preclinical rodent model. GPI may therefore contribute to life-threatening disease in humans. These data suggest that an anti-toxic vaccine against malaria might be feasible and that synthetic fragments of the *P. falciparum* GPI may be developed further to that end. □

Methods

Protein/glycan conjugation

Synthetic GPI glycan 1 was reacted with a tenfold molar excess of Traut's reagent (2-iminothiolane) in 60 mM triethanolamine, 7 mM potassium phosphate, 100 mM NaCl, 1 mM EDTA, pH 8.0 in the cold for 90 min under nitrogen, to introduce a sulphhydryl onto the free primary amine (ethanolamine). The sample was desalted by Biogel P4 filtration in coupling buffer at 4 °C, and the sample added to maleimide-activated KLH or OVA (Pierce) overnight. After exhaustive dialysis against water, conjugation efficiency was estimated by gas chromatography/mass spectroscopy. Samples were hydrolysed in 6 M HCl and the trimethylsilyl derivatives quantified for *myo*-inositol content by selective ion monitoring using *scyllo*-inositol as internal standard. For the generation of sham-conjugated carrier proteins, maleimide-activated KLH or OVA (Pierce) were subjected to identical procedures, except that cysteine was substituted for sulphhydryl-modified glycan.

Infections

All experiments were in accordance with local Animal Ethics Committee regulations. Young adult C57BL/6 mice from Jackson Laboratories were pre-bled and inoculated with 6.5 µg KLH-glycan (0.176 µg glycan, *n* = 16) or KLH-cysteine (sham-immunized, *n* = 24) emulsified in Freund's complete adjuvant, and boosted with equal amounts of immunogen in incomplete Freund's adjuvant. After two boosts, mice were rested and injected intraperitoneally with 1×10^6 erythrocytes infected with *P. berghei* ANKA. Naive mice (*n* = 12) served as unimmunized controls. Parasitemia levels were assessed from Giemsa-stained thin films. Mortality was checked twice daily. Mice were judged as developing cerebral malaria if displaying neurological signs such as loss of reflex or ataxia, or dying between days 5 and 12 after infection with relatively low parasitaemia levels. Differences in survival curves of *P. berghei*-infected mice across this time period were assessed by Cox–Mantel log rank transformation on Kaplan–Meier plots. Deaths from day 12 onwards were associated with high parasitaemia, lower rates of cerebral vascular occlusion, and anaemia as determined by haemocytometer counts.

Pathology

For histological analysis of cerebral pathology, brains were taken into 10% neutral-buffered formalin, sectioned (5 µm), and stained with haematoxylin and eosin. In other experiments, groups of six naive, sham-immunized and KLH-glycan-immunized mice were challenged as above. All mice were killed at day 6, along with age/sex-matched uninfected controls, their serum collected for determination of pH, and lungs removed. The wet weight was determined immediately after removal of the organ, and the dry weight after overnight incubation at 80 °C²³. Brains were taken for histological examination as above.

TNF output

Mycoplasma-free *P. falciparum* schizonts (3D7 strain) were prepared by gelatin flotation followed either by extraction with sample buffer (for SDS–polyacrylamide gel electrophoresis and western blots) or by saponin lysis and three washes in isotonic buffer. Parasites were taken up by sonication in complete medium, and aliquots of 5×10^6 cell equivalents in 100-µl volumes were pre-incubated for 1 h with the indicated concentration of test or control sera, followed by addition to 4×10^5 target RAW264.7 cells for 16 h in a 96-well plate. Levels of TNF-α in culture supernatants were determined by capture enzyme-linked immunosorbent assay according to manufacturer's protocol (Pharmingen) and quantified by interpolation against recombinant protein standard curves.

Immunofluorescence

Thin films of mature *P. falciparum* cultures at 10% parasitaemia were fixed in acetone at –20 °C and exposed to test and control antisera (1/80) followed, after washing in PBS, by 1/200 dilution of fluorochrome-conjugated goat anti-mouse IgG (γ-chain specific). Slides were photographed under appropriate illumination.

Statistics

For a statistical comparison between test and control groups, we used a Student's *t*-test, except for Kaplan–Meier survival plots, which were tested by Cox–Mantel log rank transformation.

Received 21 February; accepted 21 June 2002; doi:10.1038/nature00937.

1. World Health Organization World malaria situation 1990. *World Health Stat. Q.* **45**, 257–266 (1992).
2. Schofield, L. & Hackett, F. Signal transduction in host cells by a glycosylphosphatidylinositol toxin of malaria parasites. *J. Exp. Med.* **177**, 145–153 (1993).
3. Tachado, S. D. & Schofield, L. Glycosylphosphatidylinositol toxin of *Trypanosoma brucei* regulates IL-1α and TNF-α expression in macrophages by protein tyrosine kinase mediated signal transduction. *Biochem. Biophys. Res. Commun.* **205**, 984–991 (1994).
4. Schofield, L. *et al.* Glycosylphosphatidylinositol toxin of *Plasmodium* upregulates intercellular

adhesion molecule-1, vascular cell adhesion molecule-1, and E-selectin expression in vascular endothelial cells and increases leukocyte and parasite cytoadherence via tyrosine kinase-dependent signal transduction. *J. Immunol.* **156**, 1886–1896 (1996).

5. Tachado, S. D. *et al.* Glycosylphosphatidylinositol toxin of plasmodium induces nitric oxide synthase expression in macrophages and vascular endothelial cells by a protein tyrosine kinase-dependent and protein kinase C-dependent signalling pathway. *J. Immunol.* **156**, 1897–1907 (1996).
6. Tachado, S. D. *et al.* Signal transduction in macrophages by glycosylphosphatidylinositols of *Plasmodium*, *Trypanosoma* and *Leishmania*: activation of protein tyrosine kinases and protein kinase C by inositolglycan and diacylglycerol moieties. *Proc. Natl Acad. Sci. USA* **94**, 4022–4027 (1997).
7. Grau, G. E., Taylor, T. E., Molyneux, M. E., Wirima, J. J. & Vassalli, P. Tumor necrosis factor and disease severity in children with *falciparum* malaria. *N. Engl. J. Med.* **320**, 1586–1591 (1989).
8. Turner, G. D. H. *et al.* An immunohistochemical study of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. *Am. J. Pathol.* **145**, 1057–1069 (1994).
9. Magez, S. *et al.* The glycosyl-inositol-phosphate and dimyristoylglycerol moieties of the glycosylphosphatidylinositol anchor of the trypanosome variant-specific surface glycoprotein are distinct macrophage-activating factors. *J. Immunol.* **160**, 1949–1956 (1998).
10. Almeida, I. C. *et al.* Highly purified glycosylphosphatidylinositols from *Trypanosoma cruzi* are potent proinflammatory agents. *EMBO J.* **19**, 1476–1485 (2000).
11. Golgi, C. Sull' infezione malarica. *Arch. Sci. med. (Torino)* **10**, 109–135 (1886).
12. Dieckmann-Schuppert, A., Bender, S., Odenthal-Schnitler, M., Bause, E. & Schwarz, R. T. Apparent lack of N-glycosylation in the asexual intraerythrocytic stage of *Plasmodium falciparum*. *Eur. J. Biochem.* **205**, 815–825 (1992).
13. Dieckmann-Schuppert, A., Bause, E. & Schwarz, R. T. Studies on O-glycans of *Plasmodium falciparum*-infected human erythrocytes: evidence for O-GlcNAc and O-GlcNAc-transferase in malaria parasites. *Eur. J. Biochem.* **216**, 779–788 (1993).
14. Berhe, S., Schofield, L., Schwarz, R. T. & Gerold, P. Conservation of structure among glycosylphosphatidylinositol toxins from different geographic isolates of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **103**, 273–278 (1999).
15. Gowda, D. C., Gupta, P. & Davidson, E. A. Glycosylphosphatidylinositol anchors represent the major carbohydrate modification in proteins of intraerythrocytic stage *Plasmodium falciparum*. *J. Biol. Chem.* **272**, 6428–6439 (1997).
16. Gerold, P., Dieckmann-Schuppert, A. & Schwarz, R. T. Glycosylphosphatidylinositols synthesized by asexual erythrocytic stages of the malarial parasite, *Plasmodium falciparum*. Candidates for plasmoidal glycosylphosphatidylinositol membrane anchor precursors and pathogenicity factors. *J. Biol. Chem.* **269**, 2597–2606 (1994).
17. McConville, M. J. & Ferguson, M. A. J. The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem. J.* **294**, 305–324 (1993).
18. White, N. J. & Ho, M. The pathophysiology of malaria. *Adv. Parasitol.* **31**, 83–173 (1992).
19. Grau, G. E. *et al.* Tumor necrosis factor (cachectin) as an essential mediator in murine cerebral malaria. *Science* **237**, 1210–1212 (1987).
20. Grau, G. E. *et al.* Monoclonal antibody against interferon γ can prevent experimental cerebral malaria and its associated overproduction of tumour necrosis factor. *Proc. Natl Acad. Sci. USA* **86**, 5572–5574 (1989).
21. Grau, G. E. *et al.* Late administration of monoclonal antibody to leukocyte function-antigen 1 abrogates incipient murine cerebral malaria. *Eur. J. Immunol.* **21**, 2265–2267 (1991).
22. Jennings, V. M., Actor, J. K., Lal, A. A. & Hunter, R. L. Cytokine profile suggesting that murine cerebral malaria is an encephalitis. *Infect. Immun.* **65**, 4883–4887 (1997).
23. Chang, W. L. *et al.* CD8(+)–T-cell depletion ameliorates circulatory shock in *Plasmodium berghei*-infected mice. *Infect. Immun.* **69**, 7341–7348 (2001).
24. Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. The pathogenic basis of malaria. *Nature* **415**, 673–679 (2002).
25. de Souza, B. J. & Riley, E. M. Cerebral malaria: the contribution of studies in animal models to our understanding of immunopathogenesis. *Microbes Infect.* **4**, 291–300 (2002).
26. Christophers, S. R. The mechanism of immunity against malaria in communities living under hyperendemic conditions. *Ind. J. Med. Res.* **12**, 273–294 (1924).
27. Sinton, J. A. A summary of our present knowledge of the mechanism of immunity in malaria. *J. Malaria Inst. India* **2**, 71–83 (1939).
28. McGregor, I. A., Giles, H. M., Walters, J. H., Davies, A. H. & Pearson, F. A. Effects of heavy and repeated malarial infections on Gambian infants and children. *Br. Med. J.* **2**, 686–692 (1956).
29. Naik, R. S. *et al.* Glycosylphosphatidylinositol anchors of *Plasmodium falciparum*: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. *J. Exp. Med.* **192**, 1563–1576 (2000).
30. Schofield, F. Selective primary health care: strategies for control of disease in the developing world. XXII. Tetanus: a preventable problem. *Rev. Infect. Dis.* **8**, 144–156 (1986).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements

This work was supported by the UNDP/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases, the National Institutes of Health, the Human Frontiers of Science Program and the National Health and Medical Research Council. M.C.H. acknowledges a biotechnology training grant fellowship from the NIH. L.S. is an International Research Scholar of the Howard Hughes Medical Institute.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.S. (for biology) (e-mail: schofield@wehi.edu.au) or P.H.S. (for chemistry) (e-mail: seeberg@mit.edu).

Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck

Sharon Eden*, Rajat Rohatgi*, Alexandre V. Podtelejnikov†‡, Matthias Mann† & Marc W. Kirschner*

* Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue Boston, Massachusetts 02115, USA

† Protein Interaction Laboratory, University of Southern Denmark-Odense Campusvej 55, DK-5230 Odense M, Denmark

‡ Present address: MDS Proteomics, Staerosegaardsvej 6, DK-5230, Denmark

Rac signalling to actin—a pathway that is thought to be mediated by the protein Scar/WAVE (WASP (Wiskott–Aldrich syndrome protein)-family verprolin homologous protein)—has a principal role in cell motility. In an analogous pathway, direct interaction of Cdc42 with the related protein N-WASP stimulates actin polymerization¹. For the Rac–WAVE pathway, no such direct interaction has been identified. Here we report a mechanism by which Rac and the adapter protein Nck activate actin nucleation through WAVE1. WAVE1 exists in a heterotetrameric complex that includes orthologues of human PIR121 (p53-inducible messenger RNA with a relative molecular mass (M_r) of 140,000), Nap125 (NCK-associated protein with an M_r of 125,000) and HSPC300. Whereas recombinant WAVE1 is constitutively active, the WAVE1 complex is inactive. We therefore propose that Rac1 and Nck cause dissociation of the WAVE1 complex, which releases active WAVE1–HSPC300 and leads to actin nucleation.

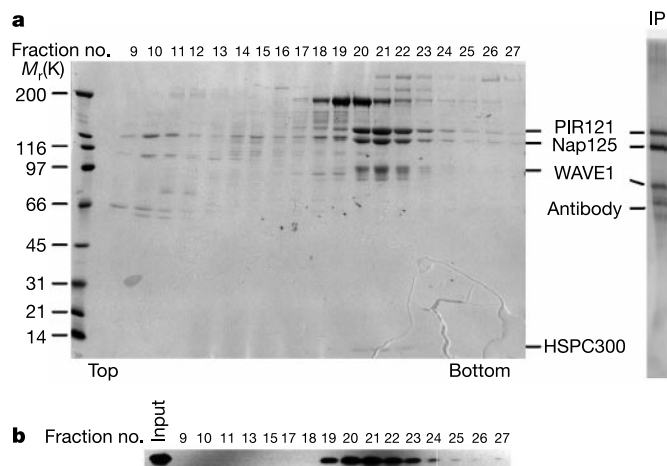


Figure 1 Purification of a WAVE1-containing complex from bovine brain. **a**, Left, Coomassie-stained 4–15% PAGE gel of fractions collected from a sucrose gradient. Right, fraction 21 was used for immunoaffinity purification (IP) using an antibody to WAVE1; the antibody heavy chain is indicated. The immunoaffinity-purified bands were analysed by MALDI. PIR121 was identified by 53 peptides covering 41% of AAD45723 (NCBI database accession number), Nap125 by 63 peptides covering 49% of NP_038464, WAVE1 by 22 peptides covering 38% of NP_003922 and HSPC300 by 12 peptides covering 45% of AAF28978. The results were confirmed by nano-electrospray tandem mass spectrometry. HSPC300 has been deposited in databases as a human gene fragment. The orthologous bovine and mouse sequences indicate a shorter coding sequence with a predicted M_r of 9K. **b**, Immunoblot using antibody to WAVE1 of the same sucrose gradient fractions as in **a**, showing an overlapping peak with the Coomassie-stained WAVE1 complex.

Members of the Rho family of small GTPases, such as Cdc42 and Rac1, and of the Src homology (SH) domain-containing SH2–SH3 adapter protein family, such as NCK, link extracellular signals and actin nucleation through pathways that include the WASP family of proteins and the actin nucleation machinery—the Arp2/3 complex¹. All WASP family members contain a conserved verprolin-homology, cofilin-homology, acidic (VCA) domain that directly binds and activates the Arp2/3 complex. The Arp2/3 complex, in turn, catalyses the nucleation of actin filaments². To prevent undesirable spontaneous actin nucleation in the absence of input signals, the activity of the WASP proteins is tightly regulated. For example, N-WASP is found predominantly in an autoinhibited conformation in which the carboxy-terminal VCA domain is occluded through interaction with the amino terminus of the protein^{3,4}. When Cdc42 binds to the Cdc42/Rac1 interactive binding (CRIB) domain of N-WASP or when NCK binds to the polyproline region of N-WASP, this autoinhibition is relieved and the VCA domain is unmasked. Phosphatidylinositol(4,5)bisphosphate (PIP₂) can further activate N-WASP in cooperation with NCK or Cdc42 by binding to a basic region of N-WASP^{5–7}.

The WAVE proteins (WAVE1, WAVE2 and WAVE3 in mammals and orthologues in *Drosophila* and *Dictyostelium*^{8–10}) are similar in structure to N-WASP⁹. They all have a C-terminal VCA domain, a polyproline region and a basic region. Unlike N-WASP, WAVE proteins do not contain a CRIB domain, and direct binding of WAVE1 to Rac1 has not been detected⁸. But much evidence suggests that WAVE1 functions downstream of Rac1. WAVE1 is translocated from the cytoplasm to membrane ruffles induced by Rac1 (ref. 8), and dominant-negative WAVE1 abolishes the formation of Rac1-dependent lamellipodia and Rac1-dependent neurite extensions⁸.

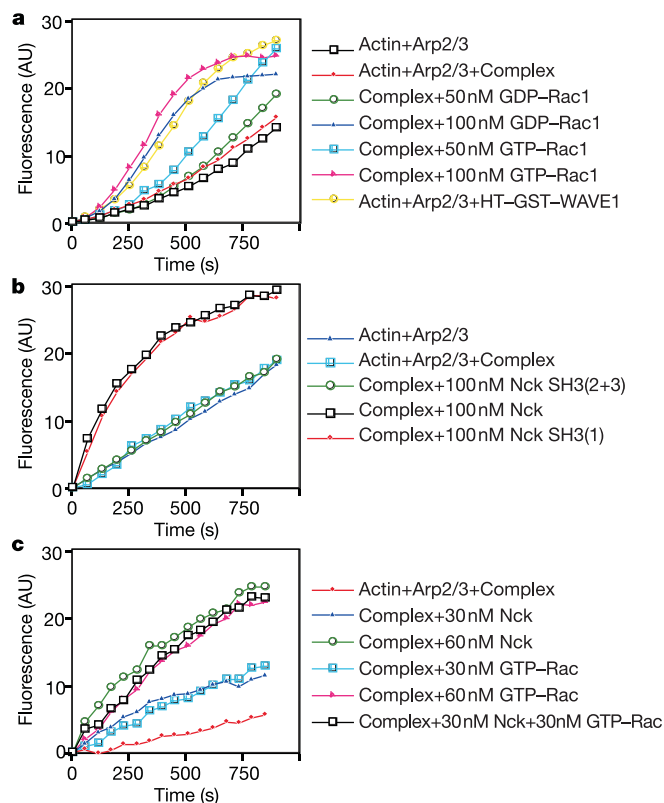


Figure 2 Rac1 or Nck induces Arp2/3 stimulation by the WAVE1 complex. Kinetics of Arp2/3-mediated actin polymerization in the absence or presence of the WAVE1 complex. GTP- or GDP-charged GST–Rac1 (**a**) or Nck fragments (**b**) or both (**c**) were added to reactions that included actin, Arp2/3 and the WAVE1 complex as indicated. Purified recombinant HT–GST–WAVE1 was added to actin and Arp2/3. AU, arbitrary units.

The mechanism of regulation of WAVE1 is likely to be fundamentally different from that of N-WASP: whereas N-WASP is auto-inhibited, recombinant WAVE1 is constitutively active in stimulating the actin nucleation activity of Arp2/3 (ref. 11). Therefore, WAVE1 activity is either inhibited *in trans* by other cellular regulators or regulated by post-translational modifications.

To identify potential *trans*-inhibitory proteins for WAVE1 activation, we searched for proteins that interact with WAVE1. We found that all of the detectable WAVE1 protein in soluble bovine brain extracts elutes in size-exclusion chromatography as a complex with an M_r of 500K (data not shown). We purified this complex to about 90% purity as judged by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining (Fig. 1a). Three proteins co-fractionated and were subsequently co-immunoprecipitated with WAVE1 (Fig. 1a).

The proteins were identified unambiguously by mass spectrometry as the bovine orthologues of the following human proteins: WAVE1; PIR121 (ref. 12), also named KIAA 0587 (ref. 13), pop (ref. 14) and CYFIP2 (ref. 15); Nap125 (ref. 16), also named NCKAP1 (ref. 17), KIAA0168 (ref. 18) and hem-2 (ref. 19); and HSPC300, which encodes a protein with an M_r of 9K (Fig. 1). The purified complex migrated with an S (Svedberg) value of 9.3 which, together with the calculated diffusion coefficient from the size-exclusion chromatography, predicts a theoretical M_r of 320K, consistent with a 1:1:1:1 stoichiometry of the proteins in the complex. Both Nap125 and PIR121 have been identified as proteins that bind directly or indirectly to NCK, profilin II and GTP-charged Rac1. Neither WAVE1 nor HSPC300 has been observed previously in screens for Rac1-, NCK- and profilin-II-interacting proteins^{14,16,20,21}.

To test whether WAVE1 activity is inhibited *in trans* by the other proteins in the complex, we tested the ability of the complex to activate Arp2/3 using a pyrene-labelled actin polymerization assay. Whereas recombinant WAVE1 activated the Arp2/3 complex, the native WAVE1 complex did not (Fig. 2a). But the activity of the complex in Arp2/3 stimulation was increased markedly by GTP γ S-charged Rac1 in a dose-dependent manner and to a lesser extent by GDP-charged Rac1. The half-time for polymerization was 3.5-fold shorter with GTP γ S-Rac1 activation and 1.5-fold shorter with GDP-Rac1, which reflects both a shorter lag time for nucleation and an increased maximal elongation rate (Fig. 2a). By contrast, GTP γ S-charged Cdc42 had no effect on the activity of the WAVE1 complex (data not shown). In the presence of GTP γ S-charged Rac1, the WAVE1 complex attained a similar activity to that of an equal concentration of purified glutathione *S*-transferase (GST)-WAVE1 (Fig. 2a), which suggests that Rac1 is fully effective in relieving the inhibition of WAVE1 in the complex.

Because both Nap125 and PIR121 associate with NCK¹⁶, we tested whether NCK can activate the WAVE1 complex. Both recombinant full-length NCK (Fig. 2b) and a recombinant fragment of NCK containing only the three SH3 domains of NCK (named SH3(1 + 2 + 3); data not shown) activated the WAVE1 complex, which shows that the SH2 domain of NCK is dispensable for activation. The association of Nap125 to NCK has been mapped to the first SH3 domain of NCK¹⁶. In agreement with this observation, the recombinant SH3(1) domain of NCK⁵ activated the complex just as effectively as did full-length NCK (Fig. 2b). By contrast, an NCK protein fragment that only contained the second and third SH3 domains⁵, SH3(2 + 3), had no activity (Fig. 2b), which indicates that SH3(1) is both necessary and sufficient for WAVE1 activation.

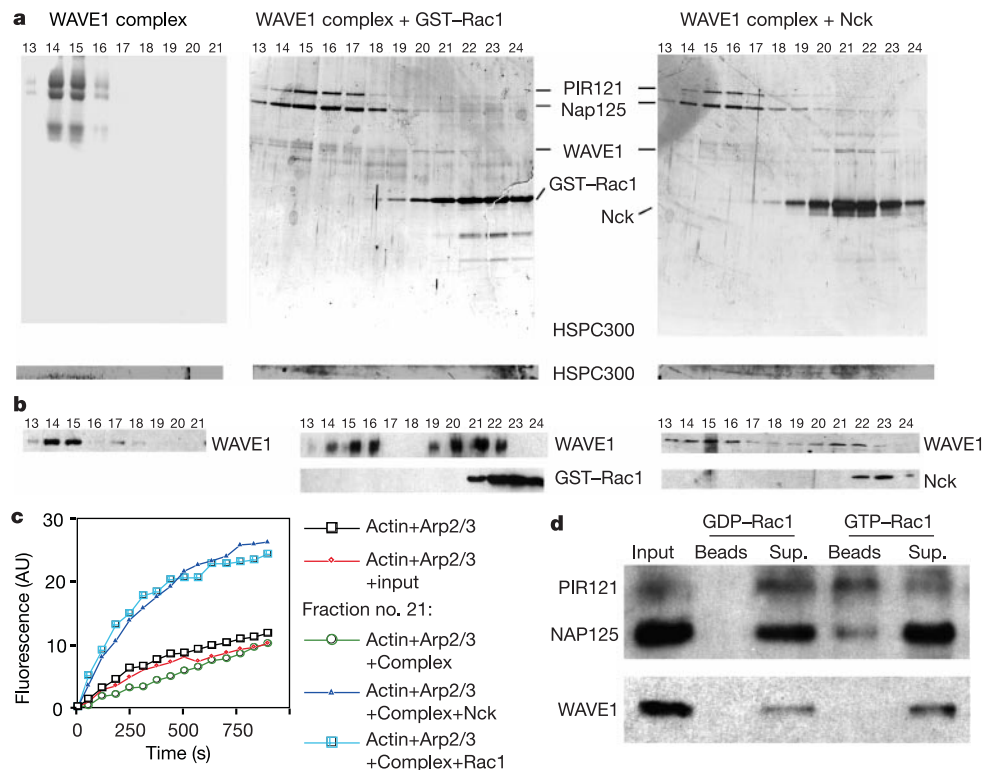


Figure 3 Disassembly of the WAVE1 complex by Rac1 and Nck enables activation. **a, b**, Purified WAVE1 complex was applied to a Superose 6 size-exclusion column in XB buffer with 10% glycerol alone (left) or in the presence of GTP γ S-charged GST-Rac1 (middle) or Nck (right). Silver-stained gels of 4-20% PAGE (**a**) and immunoblots for WAVE1 and GST-Rac1 or Nck of 10% PAGE (**b**) are shown for each sample. The background bands in the right panel of **a** are present in the original Nck bacterial

preparation. The lower panels in **a** are higher contrast exposures of the low M_r regions of the same gels. **c**, Actin polymerization assay of fraction 21 from the Superose 6 size-exclusion column. **d**, Immunoblots for PIR121, Nap125 and WAVE1 after co-precipitation assays with GST-GTP γ S-Rac1 or GST-GDP-Rac1 beads. Shown are 10% of the input and the supernatant and 100% of the beads.

Neither NCK nor GTP γ S-charged Rac1 further activated recombinant GST–WAVE1 (data not shown), which supports the idea that these activators function by relieving inhibition rather than by stimulating WAVE1 activity. Although N-WASP contains a basic region that mediates PIP₂-induced activation of N-WASP^{5,6}, the small basic region in WAVE1 does not seem to have a similar function because PIP₂ did not activate the WAVE1 complex further (data not shown). GTP γ S-charged Rac1 and NCK activated the WAVE1 complex in an additive manner (Fig. 2c).

In a simple mechanism, Rac1 and NCK could relieve the inhibition of WAVE1 in the complex by stimulating disassembly of the complex and thereby releasing active WAVE1. To test this model, we analysed the purified complex in the presence or absence of the activator proteins by size-exclusion chromatography. Whereas the intact, inactive complex eluted at an M_r of 500K, on addition of NCK or GTP γ S-charged GST–Rac1, the complex disassembled into two subcomplexes: WAVE1 and HSPC300, which co-fractionated as a complex of 180K; and PIR121 and Nap125, which co-eluted as a complex of 450K (Fig. 3a, b). Nap125 and PIR121 remained associated and could be co-immunoprecipitated (data not shown). The activator proteins (NCK and GST–Rac1) did not co-elute with PIR121 and Nap125, which possibly reflects a weakly bound complex that dissociated on passage through the column.

To determine whether disassembly of the complex contributes to activation of WAVE1, we tested the size-exclusion fractions with either NCK or GST–Rac1 for Arp2/3-dependent actin nucleation. The fractions containing PIR121 and Nap125 showed no activity (data not shown), whereas the WAVE1- and HSPC300-containing fractions (fraction 21) stimulated actin nucleation (Fig. 3c). These results indicate that NCK or Rac1 may cause disassembly of the complex and that this dissociation may activate WAVE1. To clarify further which of the proteins in the WAVE1 complex might associate with Rac1, we incubated beads coated with GTP γ S-charged GST–Rac1 with bovine brain extract. Nap125 and PIR121, but not WAVE1, associated with the beads, but showed a much lower affinity for the GDP-charged GST–Rac1 that was used as a control (Fig. 3d).

In summary, although WAVE1 had been implicated as the downstream target of Rac1, no regulatory linkage had been found previously. Consequently, the important Rac1-dependent pathway for actin nucleation has not been described. Similarly, although the association of NCK and Rac1 with NAP125 and PIR121 has been observed in several screens^{14,16,20,21}, their role as regulators of actin nucleation has not been shown. Our results indicate that WAVE1, like N-WASP and WASP, mediates signals from NCK and the Rho GTPases. The activation mechanisms of WAVE1 and N-WASP are very different: N-WASP is autoinhibited, whereas WAVE1 is *trans*-inhibited. The action of Rac1 and NCK is to disassemble the *trans*-inhibited WAVE1 complex, which releases the active WAVE1 protein in association with HSPC300. Consistent with this model, Rac1 and WAVE1 do not colocalize in the lamellipodium: WAVE1 is localized at the extreme edge of the lamellipodium, whereas Rac1 is distributed diffusely over the lamellipodium^{22,23}. Rac1 might remain bound to the NAP125 and PIR121 dissociated from the WAVE1 complex, but this has not been tested *in vivo*. Preliminary observations indicate that there is a similar regulation by *trans*-inhibition of WAVE2 in HeLa cells (H. Ho, R.R. and M.W.K., unpublished data).

Although the predominant regulation of WAVE1 activity described here is relief of *trans*-inhibition, an additional positive regulation by proteins that bind WAVE1 directly in an activator-independent manner is also possible. For example, IRSp53 has been reported to bind WAVE2 directly and enhance activation of Arp2/3 by recombinant WAVE2 (ref. 24). Preliminary data show that HSPC300, which remains associated with WAVE1 after activation, may also have a stimulating function on actin polymerization. The activation and dissociation of the WAVE1 complex process releases

a subcomplex of NAP125 and PIR121, and this subcomplex may be free to interact with other cellular components. In this way, a Rac1 or NCK signal might potentially coordinate several cellular processes—similar to pathways that are activated by the α - and the $\beta\gamma$ -subunits in heterotrimeric G-protein signalling.

Note added in proof: We have recently detected a fifth protein in the complex that co-migrates with WAVE1 on SDS–PAGE. We have identified this protein by mass spectrometry as the bovine orthologue of the human Abi2 (Abl Interactor 2) (NCBI database accession number NP_005750) (S.E., H. Ho, H. Steen and S. P. Gygi, unpublished work). Abi2 is highly homologous to hNap1 binding protein, previously described as a protein interacting with Nap125. Indeed, Abi2 appears to remain associated with the Nap125– Pir121 sub-complex upon dissociation of the WAVE1 complex (see Fig. 3). □

Methods

Purification of a WAVE1-containing complex

Throughout the purification, the WAVE1 complex was followed by immunoblotting with an antibody to WAVE1. All chromatography media described here were purchased from Pharmacia. Four bovine brains were cleaned and homogenized in XB (20 mM HEPES, pH 7.5, 100 mM KCl, 1 mM MgCl₂, 0.1 mM EDTA). The homogenate was centrifuged for 30 min at 14,000g, followed by a 30-min centrifugation at 225,000g. The supernatant was brought to 0.8 M ammonium sulphate, centrifuged at 10,000g for 10 min, and then incubated in batch with Butyl Sepharose Fast Flow for 2 h at 4 °C. After separating the flow-through from the beads, the beads were washed with five volumes of 50 mM potassium phosphate, pH 7.3, 0.4 mM EDTA containing 0.8 M ammonium sulphate, and the bound proteins were eluted in the same buffer containing 0.4 M ammonium sulphate. We concentrated the eluate (1.8 g of total protein) by precipitation with 65% ammonium sulphate. Alternatively, the brain homogenate was brought to 65% ammonium sulphate and the butyl-Sepharose binding step was omitted. The precipitate was dissolved in the same buffer, dialysed into 20 mM potassium PIPES, pH 6.8, 25 mM KCl, 1 mM MgCl₂, 0.1 mM EDTA and 1 mM dithiothreitol (DTT), loaded onto a 50-ml Resource S column and eluted with a 25–400 mM KCl gradient developed over 15 column volumes. The WAVE1 peak fractions were dialysed into the same buffer, loaded onto a Mono S column and eluted under the same conditions.

The WAVE1-containing fractions were pooled, concentrated by ultrafiltration in a 3000 pressure cell (Amicon), and loaded onto a 330-ml prep grade Superose 6 column equilibrated in 20 mM potassium-Tris, pH 8.0, 100 mM KCl, 1 mM DTT. The WAVE1 complex eluted with an apparent Stokes radius of 8 nm. The peak fractions were pooled, loaded onto a MonoQ column, and eluted with a 100–500 mM KCl gradient developed over 10 column volumes. We pooled the complex-containing fractions and loaded them onto a 5–20% (w/v) sucrose gradient poured in 20 mM HEPES, pH 7.5, 150 mM NaCl and 1 mM DTT. The gradients were centrifuged for 24 h at 178,000 g average in a SW41 rotor, and fractions were collected from the top using a gradient fractionator (ISCO). The fractions from the gradient containing the peak WAVE1 immunoreactivity were used for immunoaffinity purification. Affinity-purified antibody to WAVE1 or rabbit γ -globulin (2 mg) was crosslinked to protein A beads by dimethyl pimelimidate. The beads were incubated with the WAVE1-containing fraction for 3 h in 4 °C, washed with XB supplemented with 500 mM NaCl and eluted with 100 mM glycine, pH 2.0.

Protein identification by mass spectrometry

Protein bands separated by one-dimensional PAGE and visualized by Coomassie blue staining were excised and digested in-gel with trypsin as described²⁵. High mass accuracy matrix-assisted laser-desorption/ionization (MALDI) peptide mapping was carried out on Reflex III instrument (Bruker Daltonics). We carried out nanoelectrospray peptide sequencing²⁶ on a quadrupole time-of-flight mass spectrometer (QSTAR, PE-Sciex) using the MDS-Proteomics nanoelectrospray ion source (MDS-Proteomics). Mass spectra were searched in non-redundant protein sequence databases and expressed sequence tag databases using the PepSea search engine (MDS-Proteomics).

Protein co-precipitation

Bacterially expressed GST–Rac1 bound to glutathione beads (2 mg ml^{−1}) was charged with GDP or GTP- γ S as described²⁷ and incubated for 1 h in 4 °C with bovine brain high-speed supernatant supplemented with 1% Triton X-100. We washed the reactions three times in XB plus 500 mM KCl and twice in XB. The bound proteins were then eluted with SDS loading buffer.

Protein purification

Recombinant human Rac1 and NCK SH3(1) were expressed in *Escherichia coli* as GST fusion proteins; cleaved NCK, NCK SH3(1 + 2 + 3), NCK SH3(2 + 3) and GST–NCK SH3(2 + 3) were a kind gift from B. Mayer. Rac1 proteins were charged with GDT, GTP or GTP- γ S as described²⁷. We purified human WAVE1 from SF9 cells as a His–GST tag fusion protein. Arp2/3 complex was purified from bovine brain extract on butyl-Sepharose and DEAE Sepharose, followed by affinity purification on GST–VCA agarose and elution with 100 mM MgCl₂ (ref. 28). GST–VCA protein was expressed in *E. coli* and bound to glutathione agarose beads.

Actin polymerization assay

We prepared actin from rabbit muscle as described²⁹. Pyrene-labelled actin (15% final labelling) at 1 μ M was added to 80 nM purified His-tag-GST-WAVE1 or WAVE1-containing complex, 30 nM purified Arp2/3 complex in 20 mM HEPES, pH 7.5, 100 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA and 0.5 mM ATP. Bacterially expressed GST-Rac1 or NCK was added to a final concentration of 30 mM, 50 mM or 100 mM. The fluorescence was measured for 15 min at 21 °C using a Carry Eclipse instrument (Varian).

Antibodies

All the antibodies were generated in rabbit and affinity-purified as described³⁰. Antibody to WAVE1 was generated using amino acids 436–560 expressed in *E. coli* as an immunogen. Antibody to Nap125 was generated against the peptide sequence CHAVYKQSVTSSA. Antibody to PIR121 was generated against the peptide sequence CNEVFALNKYMKSVETDSST. Antibody to Rac1 was from Transduction Laboratories and antibody to NCK was from Neo Markers.

Received 22 February; accepted 24 May 2002; doi:10.1038/nature00859.

- Higgs, H. N. & Pollard, T. D. Regulation of actin filament network formation through ARP2/3 complex: activation by a diverse array of proteins. *Annu. Rev. Biochem.* **70**, 649–676 (2001).
- Mullins, R. D. How WASP-family proteins and the Arp2/3 complex convert intracellular signals into cytoskeletal structures. *Curr. Opin. Cell Biol.* **12**, 91–96 (2000).
- Miki, H., Sasaki, T., Takai, Y. & Takenawa, T. Induction of filopodium formation by a WASP-related actin-depolymerizing protein N-WASP. *Nature* **391**, 93–96 (1998).
- Kim, A. S., Kakalis, L. T., Abdul-Manan, N., Liu, G. A. & Rosen, M. K. Autoinhibition and activation mechanisms of the Wiskott–Aldrich syndrome protein. *Nature* **404**, 151–158 (2000).
- Rohatgi, R., Nollau, P., Ho, H. Y., Kirschner, M. W. & Mayer, B. J. Nck and phosphatidylinositol 4,5-bisphosphate synergistically activate actin polymerization through the N-WASP–Arp2/3 pathway. *J. Biol. Chem.* **276**, 26448–26452 (2001).
- Rohatgi, R., Ho, H. Y. & Kirschner, M. W. Mechanism of N-WASP activation by CDC42 and phosphatidylinositol 4,5-bisphosphate. *J. Cell Biol.* **150**, 1299–1310 (2000).
- Higgs, H. N. & Pollard, T. D. Activation by Cdc42 and PIP₂ of Wiskott–Aldrich syndrome protein (WASP) stimulates actin nucleation by Arp2/3 complex. *J. Cell Biol.* **150**, 1311–1320 (2000).
- Miki, H., Suetsugu, S. & Takenawa, T. WAVE, a novel WASP-family protein involved in actin reorganization induced by Rac. *EMBO J.* **17**, 6932–6941 (1998).
- Suetsugu, S., Miki, H. & Takenawa, T. Identification of two human WAVE/SCAR homologues as general actin regulatory molecules which associate with the Arp2/3 complex. *Biochem. Biophys. Res. Commun.* **260**, 296–302 (1999).
- Bear, J. E., Rawls, J. F. & Saxe, C. L. SCAR, a WASP-related protein, isolated as a suppressor of receptor defects in late *Dictyostelium* development. *J. Cell Biol.* **142**, 1325–1335 (1998).
- Machesky, L. M. *et al.* Scar, a WASP-related protein, activates nucleation of actin filaments by the Arp2/3 complex. *Proc. Natl Acad. Sci. USA* **96**, 3739–3744 (1999).
- Saller, E. *et al.* Increased apoptosis induction by 121F mutant p53. *EMBO J.* **18**, 4424–4437 (1999).
- Nagase, T. *et al.* Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins *in vitro*. *DNA Res.* **5**, 31–39 (1998).
- Witke, W. *et al.* In mouse brain profilin I and profilin II associate with regulators of the endocytic pathway and actin assembly. *EMBO J.* **17**, 967–976 (1998).
- Schenck, A., Bardoni, B., Moro, A., Bagni, C. & Mandel, J. L. A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related proteins FXR1P and FXR2P. *Proc. Natl Acad. Sci. USA* **98**, 8844–8849 (2001).
- Kitamura, T. *et al.* Molecular cloning of p125Nap1, a protein that associates with an SH3 domain of Nck. *Biochem. Biophys. Res. Commun.* **219**, 509–514 (1996).
- Suzuki, T. *et al.* Molecular cloning of a novel apoptosis-related gene, human Nap1 (NCKAP1), and its possible relation to Alzheimer disease. *Genomics* **63**, 246–254 (2000).
- Nagase, T., Seki, N., Ishikawa, K., Tanaka, A. & Nomura, N. Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (K1AA0161–K1AA0200) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res.* **3**, 17–24 (1996).
- Baumgartner, S. *et al.* The HEM proteins: a novel family of tissue-specific transmembrane proteins expressed from invertebrates through mammals with an essential function in oogenesis. *J. Mol. Biol.* **251**, 41–49 (1995).
- Kobayashi, K. *et al.* p140Sra-1 (specifically Rac1-associated protein) is a novel specific target for Rac1 small GTPase. *J. Biol. Chem.* **273**, 291–295 (1998).
- Kitamura, Y. *et al.* Interaction of Nck-associated protein 1 with activated GTP-binding protein Rac. *Biochem. J.* **322**, 873–878 (1997).
- Hahne, P., Sechi, A., Benesch, S. & Small, J. V. Scar/WAVE is localised at the tips of protruding lamellipodia in living cells. *FEBS Lett.* **492**, 215–220 (2001).
- Nakagawa, H. *et al.* N-WASP, WAVE and Mena play different roles in the organization of actin cytoskeleton in lamellipodia. *J. Cell Sci.* **114**, 1555–1565 (2001).
- Miki, H., Yamaguchi, H., Suetsugu, S. & Takenawa, T. IRSp53 is an essential intermediate between Rac and WAVE in the regulation of membrane ruffling. *Nature* **408**, 732–735 (2000).
- Shevchenko, A. *et al.* Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl Acad. Sci. USA* **93**, 14440–14445 (1996).
- Wilm, M. *et al.* Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466–469 (1996).
- Ma, L., Rohatgi, R. & Kirschner, M. W. The Arp2/3 complex mediates actin polymerization induced by the small GTP-binding protein Cdc42. *Proc. Natl Acad. Sci. USA* **95**, 15362–15367 (1998).
- Egile, C. *et al.* Activation of the CDC42 effector N-WASP by the Shigella flexneri LcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility. *J. Cell Biol.* **146**, 1319–1332 (1999).
- Pardee, J. D. & Spudis, J. A. Purification of muscle actin. *Methods Enzymol.* **85**, 164–181 (1982).
- Harlow, E. & Lane, D. *Antibodies: A Laboratory Manual* 283–318 (Cold Spring Harbor Press, Cold Spring Harbor, 1988).

Acknowledgements

We thank H. Ho for HT-GST-WAVE1 baculovirus; B. Mayer for Nck constructs and proteins; R. Iggo for antibody to PIR121; J. Peterson and N. Ayad for discussions, technical help and critically reading the manuscript; and L. Ma for initial observations. This work was supported by the NIH (M.W.K.) the European Molecular Biology Organization (S.E.), the Human Frontier Science Program Organization (S.E.), the Danish National Research Foundation (M.M. and A.V.P.) and the Medical Sciences Training Program (MSTP) at Harvard Medical School (R.R.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.W.K. (e-mail: marc@hms.harvard.edu).

Modulation of an RNA-binding protein by abscisic-acid-activated protein kinase

Jiaxu Li^{†‡}, Toshinori Kinoshita[‡], Sona Pandey^{*}, Carl K.-Y. Ng[†], Steven P. Gygi[§], Ken-ichiro Shimazaki[‡] & Sarah M. Assmann^{*}

^{*} Biology Department, The Pennsylvania State University, 208 Mueller Laboratory, University Park, Pennsylvania 16802, USA

[‡] Department of Biology, Faculty of Science, Kyushu University, Ropponmatsu, Fukuoka 810-8560, Japan

[§] Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

Protein kinases are involved in stress signalling in both plant and animal systems. The hormone abscisic acid mediates the responses of plants to stresses such as drought, salinity and cold. Absciscic-acid-activated protein kinase (AAPK)—found in guard cells, which control stomatal pores—has been shown to regulate plasma membrane ion channels¹. Here we show that AAPK-interacting protein 1 (AKIP1), with sequence homology to heterogeneous nuclear RNA-binding protein A/B, is a substrate of AAPK. AAPK-dependent phosphorylation is required for the interaction of AKIP1 with messenger RNA that encodes dehydrin, a protein implicated in cell protection under stress conditions. AAPK and AKIP1 are present in the guard-cell nucleus, and *in vivo* treatment of such cells with abscisic acid enhances the partitioning of AKIP1 into subnuclear foci which are reminiscent of nuclear speckles. These results show that phosphorylation-regulated RNA target discrimination by heterogeneous nuclear RNA-binding proteins² may be a general phenomenon in eukaryotes, and implicate a plant hormone in the regulation of protein dynamics during rapid subnuclear reorganization.

Guard cells regulate the apertures of microscopic stomatal pores on the leaf epidermis through which plants take up CO₂ and lose O₂ and water vapour. Appropriate regulation of stomatal aperture in response to environmental conditions, achieved by osmotic swelling and shrinking of guard cells, is vital for plant productivity and drought resistance^{3,4}. We previously used *de novo* sequencing by tandem mass spectrometry to obtain peptide sequence information that allowed us to clone the complementary DNA encoding AAPK from the fava bean, *Vicia faba*¹. Expression of a dominant negative form of AAPK in guard cells prevented activation by abscisic acid (ABA) of anion channels and stomatal closure, implicating AAPK in

[†]Present address: Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.

rapid ABA signalling events in this specialized cell type¹. To identify AAPK interaction partners, we screened a *V. faba* guard-cell cDNA expression library with radiolabelled AAPK, and thereby identified AAPK-interacting protein 1 (AKIP1). Sequence comparison (Fig. 1) indicates that AKIP1 is most similar to one of the types of protein that bind single-stranded RNA—namely, heterogeneous nuclear RNA-binding protein A/B (hnRNP A/B). Such hnRNPs comprise a varied class of proteins involved in transcriptional, post-transcriptional and translational control of gene expression^{5,6}. In particular, hnRNPs of the A/B class are involved in alternative pre-mRNA splicing, telomere biogenesis, and mRNA transport⁶.

Many eukaryotic proteins that bind single-stranded RNA contain one or more RNA recognition motifs (RRMs), consisting of a loosely conserved region of about 90 amino acids⁵. AKIP1 has a RRM between amino-acid residues 151 and 228 that is 33.3% identical at the amino-acid level to the RRM of human A/B hnRNPs⁷. There are also two highly conserved segments of amino-acid residues within RRM. The first is a hydrophobic segment of six residues called RNP-2, and the second is a segment of eight residues called RNP-1 (ref. 5). The RNP2 and RNP1 of AKIP1 are 66% and 75% identical at the amino-acid level to those of human hnRNP A/B. RNA-binding proteins also usually have auxiliary domains rich in particular amino acids. Glutamic acid

comprises 78.7% of the amino-acid residues between positions 59 and 91 of AKIP1 (Fig. 1).

Yeast two-hybrid analysis provided an *in vivo* confirmation of the AAPK–AKIP1 interaction, and indicated that the amino-terminal region of AAPK, including the entire catalytic domain, interacts with AKIP1 (Fig. 2). Ribonucleotide homopolymer binding assays using AKIP1 purified from a yeast expression system confirmed AKIP1 as an RNA-binding protein (data not shown).

Physiologically relevant interacting proteins should exhibit at least overlapping cellular localization. AAPK–GFP (green fluorescent protein) showed diffuse localization in both the cytoplasm and the nucleus (Fig. 3a), as confirmed by confocal optical sectioning and by fluorescence microscopy on isolated nuclei (data not shown). The bipartite distribution of AAPK in guard cells suggests that AAPK plays many regulatory roles, presumably by interacting with locale-specific downstream elements. AKIP1–GFP fusion protein showed nuclear localization (Fig. 3), characterized initially by a primarily diffuse nuclear labelling pattern (Fig. 3c). ABA treatment caused a rapid and notable increase in the number of AKIP1–GFP-labelled foci, indicating that ABA induces relocalization of AKIP1 into, and/or retention of AKIP1 in, nuclear speckles⁸ (Fig. 3d–f). Neither AAPK–GFP nor GFP showed this response (data not shown). To our knowledge, a hormonally induced partitioning of an hnRNP into nuclear speckles has not been previously reported⁸.

To determine whether AKIP1 is a substrate for AAPK, we performed *in vitro* and *in vivo* phosphorylation assays. Recombinant AKIP1 can be phosphorylated *in vitro* by AAPK isolated from guard cells treated with ABA (active AAPK) but not by AAPK from untreated guard cells (inactive AAPK) (Fig. 4a), consistent with previous observations of the ABA-dependence of AAPK activation^{9,10}. ABA-dependent phosphorylation of AKIP1 occurs *in vivo* in guard cells: phosphorylated AKIP1 was recovered by immunoprecipitation from intact guard-cell protoplasts labelled *in vivo* with ³²P-orthophosphate and treated for 15 min with ABA (Fig. 4b). ABA treatment did not alter AKIP1 abundance (Fig. 4c). Collectively, these results demonstrate that AKIP1 is indeed a substrate of AAPK.

We sought the RNA targets of AKIP1 in order to further assess the functional significance of AKIP1 phosphorylation by AAPK. Dehydrins are ubiquitous stress-upregulated proteins thought to improve the stability of enzymes and membranes under stress conditions^{11,12}. In addition, dehydrins have been demonstrated to be transcribed more abundantly in ABA-treated *V. faba* guard cells¹². Accordingly, we amplified by 3' end amplification (TPEA; ref. 13) guard-cell mRNA bound to AKIP1, and then performed polymerase chain reactions (PCR) using dehydrin-specific primers. As confirmed by sequence analysis of the PCR product (see also Fig. 5a), dehydrin could be amplified from the mRNA population bound to AKIP1, but only if AKIP1 had been phosphorylated by

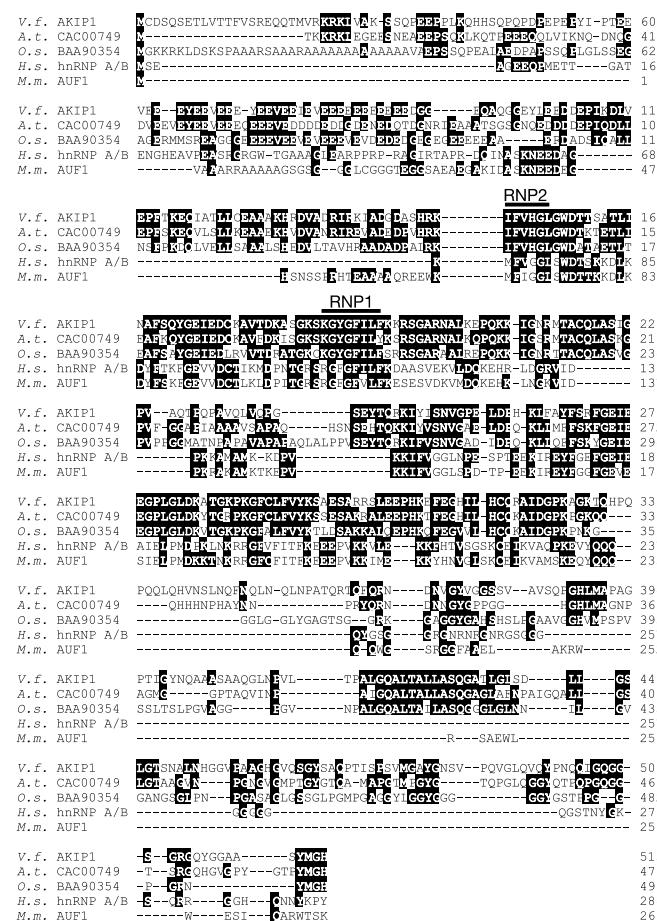


Figure 1 AKIP1 encodes a protein that binds single-stranded RNA. The deduced amino-acid sequence of AKIP1 cDNA aligned (Clustal method) with several putative single-stranded RNA-binding proteins. A.t. CAC00749 and O.s. BAA90354 (GenBank) are from *Arabidopsis thaliana* and *Oryza sativa* sequencing projects; H.s. hnRNP A/B (heterogeneous nuclear RNA-binding protein A/B) and M.m. AUF1 are from human and mouse with accession numbers S17563 and NP_031542 (GenBank), respectively. Identities are shaded in black, and gaps are indicated with dashes. The numbers on the right refer to amino-acid positions. RNP1 and RNP2 consensus sequences are indicated.

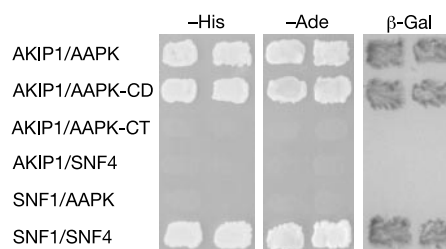


Figure 2 AKIP1 specifically interacts with AAPK in the yeast two-hybrid system. Positive interactions are indicated by growth on media lacking histidine (–His; left), or lacking adenine (–Ade; middle), and by the expression of β-galactosidase (β-Gal; right). AAPK-CD and AAPK-CT represent the N-terminal region including the entire catalytic domain of AAPK (residues 1–267 of the 349 amino acids) and the C-terminal region including the acidic region of AAPK (residues 263–349), respectively. Yeast SNF1 and SNF4 are interacting positive controls.

pretreatment with ABA-activated AAPK (Fig. 5a lane 5 versus lane 6). Thus, AKIP1 phosphorylation by AAPK increased the affinity of AKIP1 for a particular transcript.

The association of dehydrin transcript with AKIP1 was also only observed when the poly(A) RNA target population was isolated from guard cells initially treated with ABA (Fig. 5a, lanes 3 v. 6). This initial ABA treatment increased dehydrin transcript abundance (Supplementary Information), plausibly increasing the number of dehydrin transcripts bound to AKIP1 and thereby facilitating amplification.

To confirm direct interaction between AKIP1 and dehydrin transcript, we performed ultraviolet (UV) crosslinking RNA-binding analysis using 32 P-labelled dehydrin mRNA transcribed *in vitro*. AKIP1 binding of the radiolabelled dehydrin RNA was indeed observed (Fig. 5b). Consistent with Fig. 5a, this binding was dependent on AKIP1 phosphorylation by ABA-activated AAPK. Binding of labelled sense dehydrin mRNA was competed with cold dehydrin mRNA, demonstrating the specific nature of the dehydrin mRNA–AKIP1 interaction. Incubation of AKIP1 with active calcium-dependent protein kinase (CDPK), also purified from guard

cells¹⁴, did not suffice to allow dehydrin transcript association with AKIP1, indicating kinase specificity of the AKIP1 modification. Further, no RNA binding by AKIP1 was observed with dehydrin RNA transcribed in the antisense direction, nor with RNA transcribed from positive control template provided with the Riboprobe *in vitro* transcription kit (Promega), confirming that AKIP1 does not indiscriminately bind RNA species. Collectively, our results are consistent with the following sequence: (1) ABA activates AAPK; (2) AAPK phosphorylates AKIP1; (3) as a result, AKIP1 shows a phosphorylation-induced increase in affinity for dehydrin mRNA concomitant with increased localization in nuclear speckles.

The observation that phosphorylation alters the affinity of a plant hnRNP-type protein for a specific mRNA has not, to our knowledge, been made before. One plant hnRNP, MA16 from maize, has been shown to be phosphorylated by *in vivo* labelling with 32 P-orthophosphate, but phosphorylation status had no effect on subsequent ribohomopolymer affinity¹⁵. In fact, little is known about the functional roles of plant hnRNPs^{16,17} apart from assumptions based on sequence homology. One exception is UBP1 hnRNP of *Nicotiana plumbaginifolia* which has been shown to enhance the splicing of inefficiently processed introns^{16–18}. Mutations in other, non-hnRNP types of RNA-binding proteins affect flowering and hormone sensitivity^{19–23}. cDNAs encoding proteins that bind ribohomopolymers and whose transcription is upregulated in response to ABA have been identified in several species^{24–26}, but the functions of these proteins remain unknown. In contrast, neither AKIP1 nor AAPK transcript abundance is ABA-regulated (data not shown).

Abiotic stresses are established modulators of mRNA stability in plants; our data suggest the involvement of hnRNPs in this phenomenon. The ABA-regulated nuclear localization pattern of AKIP1 in distinct foci is consistent with a role for this protein in RNA metabolism, because nuclear speckles of mammalian systems are implicated in splicing component storage, prespliceosome assembly, and possibly active RNA processing^{8,26}. hnRNPs usually interact with multiple RNA partners⁶, and following whole-plant treatment with ABA, numerous transcripts show an increase in steady-state levels¹¹. In future research it will be of interest to

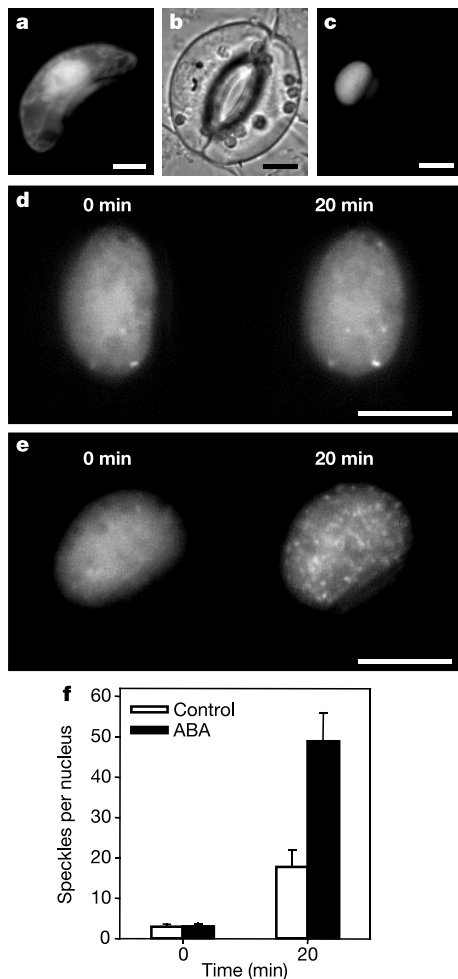


Figure 3 AKIP1 and AAPK localization in guard cells. **a**, AAPK–GFP localizes in the nucleus and cytoplasm. **b**, Bright-field image of a stoma, showing the nucleus of left guard cell with gold particles evident. **c**, Fluorescence image corresponding to **b** showing diffuse nuclear localization of AKIP1–GFP (no abscisic acid (ABA) treatment). **d**, **e**, Representative guard-cell nucleus expressing AKIP1–GFP at 0 and 20 min following incubation in **d**, control buffer ($n = 19$) or **e**, 50 μ M ABA ($n = 14$). **f**, Number of AKIP1–GFP labelled speckles following treatment with ($n = 14$) or without ($n = 19$) 50 μ M ABA. At 20 min, +ABA differs from control at $P \leq 0.05$ (analysis of variance). Scale bar, 10 μ m.

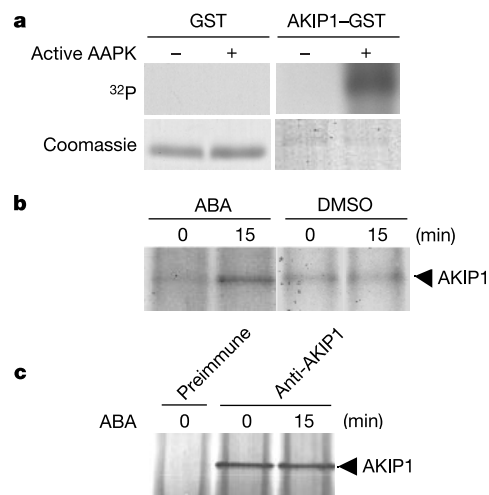


Figure 4 Phosphorylation of AKIP1 by ABA-activated AAPK regulates AKIP1 binding to dehydrin mRNA. **a**, AAPK phosphorylates recombinant AKIP1 in an ABA-dependent manner. *In vitro* phosphorylation assay was performed as described in the Methods section. GST, glutathione *S*-transferase. **b**, ABA-dependent phosphorylation of AKIP1 occurs *in vivo*. AKIP1 was immunoprecipitated from 32 P-labelled guard-cell protoplasts using AKIP1-specific antibodies. ABA was added at 25 μ M with 0.25% DMSO (dimethyl sulphoxide) as vehicle; DMSO alone had no effect. **c**, ABA treatment does not alter AKIP1 protein abundance. AKIP1 was immunoprecipitated from guard-cell protoplasts and detected by immunoblot. ABA treatment was as for **b**.

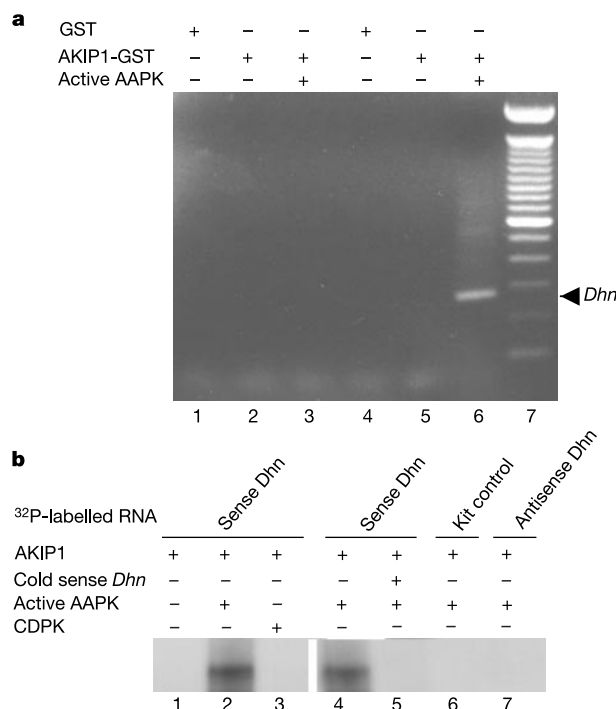


Figure 5 Interaction of AAPK-phosphorylated AKIP1 with dehydrin transcript. **a**, AAPK-phosphorylated AKIP1 binds dehydrin (*Dhn*) mRNA. RNA-binding assay was performed as described in Methods. A ~300-base-pair dehydrin PCR product was amplified from RNAs bound to AAPK-phosphorylated AKIP1 (lane 6, arrowhead). Poly(A) RNA was isolated from guard-cell protoplasts without (lanes 1–3) or with lanes 4–6 prior treatment with 25 μ M ABA. Markers of relative molecular mass are shown on the right. **b**, AAPK-phosphorylated AKIP1 specifically interacts with sense dehydrin RNA. ³²P-labelled sense dehydrin RNA

was incubated with AKIP1–GST previously treated with inactive AAPK (lane 1), active AAPK (lane 2), or active CDPK (lane 3). Alternatively, AKIP1 treated with active AAPK was incubated with ³²P-labelled sense dehydrin RNA (lane 4), or first incubated with an excess of unlabelled (competitor) sense dehydrin then with ³²P-labelled sense dehydrin RNA (lane 5), or with unrelated ³²P-labelled RNA (non-competitor) from the RNA synthesis kit (lane 6), or with ³²P-labelled antisense dehydrin RNA (*Dhn*, lane 7). Binding of ³²P-labelled RNA to AKIP1 was visualized by autoradiography.

determine the extent to which these or other transcripts comprise AKIP1 targets. Although AAPK gene expression appears to be guard-cell specific¹, AKIP1 is also expressed in leaves, stems and roots (data not shown), raising the possibility that other cell- or organ-specific kinases also differentially regulate AKIP1.

ABA has long been known as a transcriptional regulator¹¹, and as a post-translational regulator of cellular and long-distance signalling in plants^{3,4,11}. However, recent reports also show altered plant sensitivity to ABA following mutation of a double-stranded RNA-binding protein, an Sm-like protein, and an mRNA cap protein^{21–23}. The present work on the hnRNP-like protein AKIP1 complements these studies, and taken together with them, suggests a new paradigm involving ABA-regulation of post-transcriptional RNA metabolism. The present observations that ABA functions via AAPK to induce rapid changes in nuclear architecture and AKIP1 target affinity are consistent with this suggestion. □

Methods

Interaction cloning and yeast two-hybrid analysis

AAPK protein was labelled with [³⁵S]methionine using an *in vitro* transcription/translation system (Novagen), and used to probe²⁷ a *V. faba* guard-cell cDNA expression library¹ for AAPK-interacting clones. The coding sequences of AKIP1 and AAPK cDNAs were cloned into pAS2-1 and pACT2 (Clontech), respectively. Interaction of AAPK and AKIP1 was examined by yeast two-hybrid analysis using strain PJ69-4A.

AKIP1 subcellular localization

Leaves of *V. faba* were transformed by particle bombardment with AAPK–GFP or AKIP1–GFP plasmid. Expression of AAPK–GFP or AKIP1–GFP in guard cells was examined by fluorescence microscopy¹. Fluorescence images of AKIP1–GFP distribution were acquired with a CCD camera (Photometrics). The number of AKIP1–GFP labelled speckles was quantified from images taken at 0 and 20 min after incubation in 10 mM MES (2-[N-morpholino]ethanesulphonic acid), 50 mM KCl, pH 6.2, with or without 50 μ M ABA.

Production and purification of fusion protein

AKIP1 was cloned into the pESP-3 vector (Stratagene) with AKIP1 upstream of the FLAG

tag and glutathione S-transferase (GST) sequences, and was confirmed by DNA sequencing. AKIP1–GST fusion protein expressed in *Schizosaccharomyces pombe* was affinity-purified using glutathione-sepharose columns. Purified AKIP1–GST fusion protein was recognized specifically by FLAG antibody (Stratagene) using immunoblot analysis (data not shown).

In vitro phosphorylation assay

Purified AKIP1–GST or GST was subjected to *in vitro* phosphorylation assays using ABA-activated or inactive AAPK, as previously described for AAPK phosphorylation of histone⁹. Because AAPK cannot be activated *in vitro* by ABA, isolated guard-cell protoplasts were treated with ABA, and native guard-cell AAPK (or CDPK) then was gel-purified following gel renaturation^{9,14}.

In vivo phosphorylation assay

A protein of approximate relative molecular mass 70,000 ($M_r \approx 70K$) affinity-purified from *Escherichia coli* expressing an AKIP1–HIS₆ construct (Promega) was in-gel digested with trypsin. Resulting peptides were analysed by tandem mass spectrometry²⁸ on an LCQ-DECA mass spectrometer (ThermoFinnigan). AKIP1 identity was unambiguously confirmed by matching of 5 peptides using the Sequest algorithm. AKIP1 antibodies were then obtained from rabbit using recombinant AKIP1 as antigen. In immunoblots of guard-cell proteins, these antibodies specifically recognized AKIP1 protein of $M_r \approx 75K$; bands were not detected with preimmune serum (data not shown).

Guard-cell protoplast suspensions (1.0 mg protein ml⁻¹) were incubated with ³²P-orthophosphate (0.25 mCi per mg protein) under background red light for 80 min at 24 °C, then ABA was added to the suspension at 25 μ M. Immunoprecipitation and *in vivo* phosphorylation assays were performed as described²⁹.

mRNA binding assay

Approximately 3 μ g of poly(A) RNA was isolated from about 2.5×10^8 purified guard cells untreated or treated with 25 μ M ABA⁹. Purified AKIP1–GST or GST (10 μ g) was treated with active or inactive AAPK, and then incubated with GST affinity resin (Stratagene) equilibrated in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 5 mM MgCl₂ for 30 min at 4 °C. After washing 3 times to remove any unbound proteins, immobilized AKIP1–GST or immobilized GST was incubated with poly(A) RNA (~0.4 μ g) at 30 °C for 10 min. The resin was washed with the above buffer 4 times to remove unbound RNAs. Bound RNAs were then extracted from the resin as described³⁰.

A TPEA method allowing the assay of specific mRNAs was used¹³. Recovered RNAs were reverse-transcribed with an anchored oligo primer [5'-CGC TCG AGT GCA GAA TTC TTT TTT TTT TTT TTT TTT T(A,C,G)-3'] and second-strand cDNA was synthesized

using the primer [5'-GCC GCT CGA GTG CAG AT TCN NNN NCG AGA-3']. cDNAs were further amplified using the following primers: [5'-CGC TCG AGT GCA GAA TTC-3'] and [5'-GCC GCT CGA GTG CAG AT TCN NNN NCG AGA-3']. Primers were designed based on ref. 13, with modification to minimize the likelihood of inadvertently favouring the amplification of any specific known plant DNA sequence(s). The resulting cDNAs were further amplified by PCR using primers: 5'-AAC AGG GTA CGG TGG AAG TG-3' and 5'-ATC CTC CAG TAC CAG GAA GC-3', which correspond to nucleotide positions 395 to 675 of dehydrin *DHN1* (GenBank accession no. X63061)¹². The resultant PCR product was cloned into TOPO TA cloning vector (Invitrogen) and sequenced.

UV crosslinking assay

The dehydrin PCR product was subcloned into pGEM-3Z vector (Promega). After linearizing the vector with *Sma*I, dehydrin RNA was synthesized using an *in vitro* transcription system (Promega) in the presence of [α -³²P]rCTP according to Promega's protocol. Unincorporated nucleotides were removed by a NucTrap probe purification column (Stratagene). Purified AKIP1-GST or GST (2 μ g) previously treated with active AAPK or CDPK was incubated with ³²P-labelled dehydrin RNA for 30 min at room temperature. UV crosslinking and detection of RNA-bound proteins were performed as described³⁰.

Received 22 April; accepted 14 June 2002; doi:10.1038/nature00936.

- Li, J., Wang, X. Q., Watson, M. B. & Assmann, S. M. Regulation of abscisic acid-induced stomatal closure and anion channels by guard cell AAPK kinase. *Science* **287**, 300–303 (2000).
- Ostrowski, J. *et al.* Insulin alters heterogeneous nuclear ribonucleoprotein K protein binding to DNA and RNA. *Proc. Natl Acad. Sci. USA* **98**, 9044–9049 (2001).
- Schroeder, J. I., Kwak, J. M. & Allen, G. J. Guard cell abscisic acid signalling and engineering drought hardiness in plants. *Nature* **410**, 327–330 (2001).
- Assmann, S. M. & Wang, X.-Q. From milliseconds to millions of years: guard cells and environmental responses. *Curr. Opin. Plant Biol.* **4**, 421–428 (2001).
- Burd, C. G. & Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615–621 (1994).
- Krecic, A. M. & Swanson, M. S. hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11**, 363–371 (1999).
- Khan, F. A., Jaiswal, A. K. & Szer, W. Cloning and sequence analysis of a human type A/B hnRNP protein. *FEBS Lett.* **290**, 159–161 (1991).
- Misteli, T. Protein dynamics: implications for nuclear architecture and gene expression. *Science* **291**, 843–847 (2001).
- Li, J. & Assmann, S. M. An ABA-activated and calcium-independent protein kinase from guard cells of fava bean. *Plant Cell* **8**, 2359–2368 (1996).
- Mori, I. C. & Muto, S. Abscisic acid activates a 48-kilodalton protein kinase in guard cell protoplasts. *Plant Physiol.* **113**, 833–839 (1997).
- Busk, P. K. & Pages, M. Regulation of abscisic acid-induced transcription. *Plant Mol. Biol.* **37**, 425–435 (1998).
- Shen, L., Outlaw, W. H. Jr & Epstein, L. M. Expression of an mRNA with sequence similarity to pea dehydrin (*Psdhn 1*) in guard cells of *Vicia faba* in response to exogenous abscisic acid. *Physiol. Planta* **95**, 99–105 (1995).
- Dixon, A. K., Richardson, P. J., Lee, K., Carter, N. P. & Freeman, T. C. Expression profiling of single cells using 3 prime end amplification (TPEA) PCR. *Nucleic Acids Res.* **26**, 4426–4431 (1998).
- Li, J., Lee, Y. R. & Assmann, S. M. Guard cells possess a calcium-dependent protein kinase that phosphorylates the KAT1 potassium maturation. *Plant Physiol.* **116**, 785–795 (1998).
- Freire, M. A. & Pages, M. Functional characteristics of the maize RNA-binding protein MA16. *Plant Mol. Biol.* **29**, 797–807 (1995).
- Lorkovic, Z. J. & Barta, A. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Res.* **30**, 623–635 (2002).
- Lorkovic, Z. J., Wiczkorek, K. D., Lamberman, M. H. L. & Filipowicz, W. Pre-mRNA splicing in higher plants. *Trends Plant Sci.* **5**, 160–167 (2000).
- Lamberman, M. H. L. *et al.* UBPL1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* **19**, 1638–1649 (2000).
- Macknight, R. *et al.* *FCA*, a gene controlling flowering time in *Arabidopsis*, encodes a protein containing RNA-binding domains. *Cell* **89**, 737–745 (1997).
- Schomburg, F. M., Patton, D. A., Meinke, D. W. & Amasino, R. M. *FPA*, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. *Plant Cell* **13**, 1427–1436 (2001).
- Lu, C. & Fedoroff, N. A mutation in the *Arabidopsis* *HYL1* gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *Plant Cell* **12**, 2351–2366 (2000).
- Xiong, L. *et al.* Modulation of abscisic acid signal transduction and biosynthesis by an Sm-like protein in *Arabidopsis*. *Dev. Cell* **1**, 771–781 (2001).
- Hugouvieux, V., Kwak, J. M. & Schroeder, J. I. An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in *Arabidopsis*. *Cell* **106**, 477–487 (2001).
- Gómez, J. *et al.* A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. *Nature* **344**, 262–264 (1988).
- Dunn, M. A., Brown, K., Lightowler, R. & Hughes, M. A. A low-temperature-responsive gene from barley encodes a protein with single-stranded nucleic acid-binding activity which is phosphorylated *in vitro*. *Plant Mol. Biol.* **30**, 947–959 (1996).
- Melcak, I., Melcakova, S., Kopsky, V., Vecerova, J. & Raska, I. Prespliceosomal assembly on microinjected precursor mRNA takes place in nuclear speckles. *Mol. Biol. Cell* **12**, 393–406 (2001).
- Stone, J. M., Collinge, M. A., Smith, R. D., Horn, M. A. & Walker, J. C. Interaction of a protein phosphatase with an *Arabidopsis* serine-threonine receptor kinase. *Science* **266**, 793–795 (1994).
- Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).
- Kinoshita, T. & Shimazaki, K. Blue light activates the plasma membrane H⁺-ATPase by phosphorylation of C-terminus in stomatal guard cells. *EMBO J.* **18**, 5548–5558 (1999).

- Moz, Y., Silver, J. & Naveh-Man, T. Protein-RNA interactions determine the stability of the renal NaPi-2 cotransporter mRNA and its translation in hypophosphatemic rats. *J. Biol. Chem.* **274**, 25266–25272 (1999).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements

We thank S. Gilroy and E. Kunze for the use of their confocal and fluorescence microscopes; M. Guiltinan for use of the gene gun; H. Ma for providing the SNF1 and SNF4 constructs; P. James for providing the PJ69-4A strain; L. Ding for assistance with AKIP1 constructs; and P. Minnich for technical support. This research was supported by NSF (S.M.A.) and partly supported by a Grant-in-Aid for Scientific Research on Priority Areas from Ministry of Education, Sports and Culture of Japan (K.S. and T.K.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence should be addressed to J.L. (e-mail: jiaxu_li@hms.harvard.edu) or S.M.A. (e-mail: sma3@psu.edu). Requests for materials should be addressed to S.M.A. The sequence of the *AKIP1* cDNA has been deposited in GenBank, accession no. AF403292.

Transcriptional co-activator PGC-1 α drives the formation of slow-twitch muscle fibres

Jiandie Lin*, Hai Wu†, Paul T. Tarr*, Chen-Yu Zhang‡, Zhidan Wu*, Olivier Boss‡, Laura F. Michael*, Pere Puigserver*, Eiji Isotani§, Eric N. Olson†, Bradford B. Lowell‡, Rhonda Bassel-Duby|| & Bruce M. Spiegelman*

* Dana-Farber Cancer Institute and the Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Department of Molecular Biology, § Pathology, and || Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

‡ Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02115, USA

The biochemical basis for the regulation of fibre-type determination in skeletal muscle is not well understood. In addition to the expression of particular myofibrillar proteins, type I (slow-twitch) fibres are much higher in mitochondrial content and are more dependent on oxidative metabolism than type II (fast-twitch) fibres¹. We have previously identified a transcriptional co-activator, peroxisome-proliferator-activated receptor- γ co-activator-1 (PGC-1 α), which is expressed in several tissues including brown fat and skeletal muscle, and that activates mitochondrial biogenesis and oxidative metabolism^{2–4}. We show here that PGC-1 α is expressed preferentially in muscle enriched in type I fibres. When PGC-1 α is expressed at physiological levels in transgenic mice driven by a muscle creatine kinase (MCK) promoter, a fibre type conversion is observed: muscles normally rich in type II fibres are redder and activate genes of mitochondrial oxidative metabolism. Notably, putative type II muscles from PGC-1 α transgenic mice also express proteins characteristic of type I fibres, such as troponin I (slow) and myoglobin, and show a much greater resistance to electrically stimulated fatigue. Using fibre-type-specific promoters, we show in cultured muscle cells that PGC-1 α activates transcription in cooperation with Mef2 proteins and serves as a target for calcineurin signalling, which has been implicated in slow fibre gene expression. These data indicate that PGC-1 α is a principal factor regulating muscle fibre type determination.

Skeletal muscles of vertebrates contain myofibres that differ in contractile function, mitochondrial content and metabolic proper-

ties¹. Slow-twitch (type I) myofibres contain a high concentration of mitochondria and use oxidative metabolism as a chief energy source. In contrast, fast-twitch myofibres, such as type IIb fibres, generate ATP mainly through glycolysis. Muscles rich in type II fibres are more susceptible to fatigue in part because their glycolytic metabolism causes acidification of the muscle bed under repeated use. These different fibre types also synthesize distinct arrays of contractile proteins, enzymes involved in intermediary metabolism, and regulatory proteins that act together to maintain a specific fibre phenotype^{1,5}.

Adult skeletal muscle displays plasticity that allows conversion of different fibre types in response to chronic change in contractile demands^{5–8}. For example, repeated mechanical overload and endurance training increases the percentage of type I fibres and these effects can be mimicked by electrical stimulation of motor nerves or cross-reinnervation of muscles with nerves supplying slow-twitch fibres. Studies of pathways downstream of neural activity have implicated calcium signalling through calcineurin—a calcium-calmodulin (CaM)-dependent protein phosphatase—and CaM kinase in the control of type-I-fibre-specific contractile proteins^{9–12}. However, the nuclear targets ultimately responsible for the marked and stable changes in gene expression and biological function are not well understood.

We have previously shown that PGC-1 α (also known as Pgc-1 and Ppargc1) is a potent transcription co-activator for nuclear receptors and certain other transcription factors^{2,3}. PGC-1 α is expressed in skeletal muscle, and powerfully induces mitochondrial biogenesis when expressed ectopically in skeletal and cardiac myocytes^{3,13}. PGC-1 α is also critically involved in other aspects of energy metabolism, such as adaptive thermogenesis in brown fat and hepatic gluconeogenesis^{2,14}. The particularly robust role of PGC-1 α in mitochondrial biogenesis, and the fact that mitochondrial metabolism is viewed as a critical part of the muscle-fibre-type phenotype led us to examine a potential role for PGC-1 α in the control of specification of muscle fibre type. We first examined the expression of PGC-1 α in murine muscles that differ in their predominant fibre types (Fig. 1a). PGC-1 α messenger RNA expression mimics that of troponin I (slow), a classical marker of type I fibres, being very highly expressed in soleus but having much lower expression in type-II-rich muscles such as extensor digitorum longus (EDL), quadriceps, tibialis anterior, and gastrocnemius. Several PGC-1 α mRNA species are derived from the usage of alternative polyadenylation sites¹⁵. PGC-1 β ¹⁶, a recently described homologue of PGC-1 α , does not share this muscle type selectivity.

To more directly assess the role for PGC-1 α in myofibre specification, we generated transgenic mice expressing PGC-1 α under control of the MCK promoter, a promoter that is preferentially activated in type II fibres¹⁷. As shown in Fig. 1b, PGC-1 α mRNA is expressed at various levels in different leg muscles from four transgenic lines, with highest expression observed in lines 23 and the next highest in line 31. Analysis of protein levels in different isolated muscles from mice of strain 31 showed that PGC-1 α was elevated in plantaris—a muscle containing primarily type II fibres—from an undetectable level to the same level as that seen in soleus muscle, a muscle rich in type I fibres. PGC-1 α protein levels are not significantly elevated in the soleus muscle of this strain (Fig. 1c). Transgenic PGC-1 α expression induced mitochondrial genes, as evidenced by increased expression of mRNA for mitochondrial enzymes, such as cytochrome *c* oxidase (COX) II and COX IV, enzymes involved in electron transport (Fig. 1b). In addition to the activation of genes involved in oxidative energy production, PGC-1 α robustly induced the expression of genes specifically enriched in type I fibres, such as myoglobin and troponin I (slow), in both of these strains (Fig. 1b). A corresponding decrease in troponin I (fast) expression was observed in line 23, suggesting that a more complete fibre type switching occurred in the presence of higher PGC-1 α concentrations. However, this mouse strain also showed a distinct

weight loss and a loss of reproductive capacity, and was therefore not used for further functional analysis beyond characterization of gene expression patterns.

From a gross morphological perspective, most of the muscles in transgenic strain 31 showed the distinct red colour characteristic of oxidative muscle, whereas the muscles of non-transgenic littermates were paler in appearance (Fig. 2a, b). Gastrocnemius, which contains a mixture of type I and II fibres, can be readily distinguished from the dark red, type-I-enriched soleus in wild-type mice; however, these muscles were not distinguishable by colour in transgenic mice. Consistent with these morphological changes, expression of myoglobin, troponin I (slow), and cytochrome *c* was highly elevated in type II white vastus muscle of transgenic mice (Fig. 2c). Metachromatic staining can distinguish between type I, IIa and IIb fibres¹⁸. Histology with metachromatic staining and immunohistochemistry with antibodies to the slow fibre myosin revealed the presence of type I and IIa fibres in plantaris (about 10% and 20%, respectively), a muscle group containing almost exclusively type IIb fibres in non-transgenic mice (Fig. 2d). These results demonstrate that increased expression of PGC-1 α in type II myofibres activates a coordinated programme that is characteristic of type I fibres.

A principal functional hallmark of type I fibres is their fatigue-resistance and ability to sustain a prolonged duration of muscle activity. To determine whether transgenic muscle is more resistant to fatigue, EDL muscles were isolated and subjected to continuous electrical stimulation that mimics muscle contraction during exercise. In a standard assay, we measured the time required for the muscle contraction force to fall below 30% of the starting level (the 'fatigue index'¹⁹). As shown in Fig. 2e, transgenic muscles were able to sustain a much longer duration of continuous contraction than

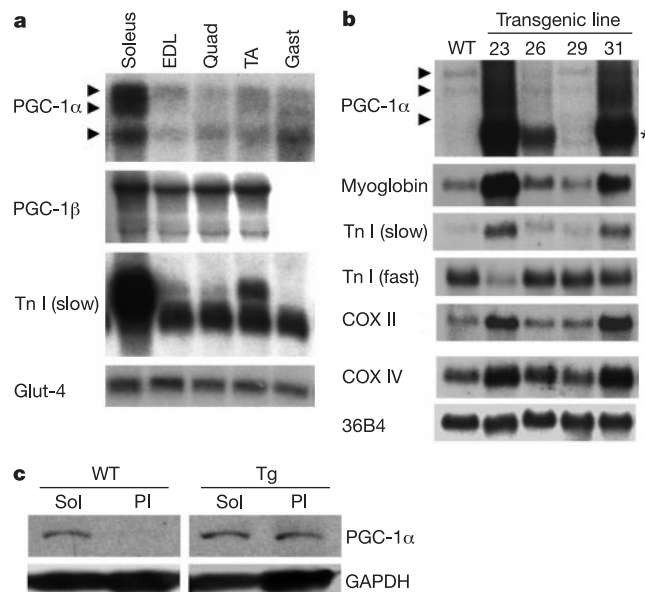


Figure 1 Analysis of gene expression in wild-type (WT) and transgenic (Tg) muscle. **a**, Expression of *PGC-1 α* and other genes in various dissected wild-type muscles. EDL, extensor digitorum longus; quad, quadriceps; TA, tibialis anterior; gast, gastrocnemius; Tn I (slow), troponin I (slow). Arrowheads indicate *PGC-1* mRNA species. **b**, RNAs were isolated from gastrocnemius and subjected to hybridization analysis with probes for fibre-specific contractile proteins and mitochondrial enzymes. Hybridization with ribosomal protein 36B4 probes served as a loading control. Asterisk indicates transgenic *PGC-1* mRNA. COX II and IV, cytochrome *c* oxidase II and IV. **c**, Immunoblot analysis of *PGC-1 α* expression in type-I-rich soleus (sol) and type-IIb-rich plantaris (pl), with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as a loading control.

wild-type controls, indicating that contractile properties of transgenic muscles underwent functional changes characteristic of oxidative fibres. Importantly, the fatigue index of PGC-1 α transgenic muscle is the highest among several mouse models for fibre-type conversion, including calcineurin and CaM kinase transgenics and exercise (data not shown).

Calcium signalling through calcineurin is a chief regulatory pathway of type I fibre-selective gene expression¹². We assayed the ability of PGC-1 α to stimulate transcription from the myoglobin and troponin I (slow) promoters, with or without the addition of calcineurin, in cultured C2C12 muscle cells. In the absence of calcineurin, PGC-1 α activated these promoters poorly (Fig. 3a); however, PGC-1 α increased the activity of the myoglobin and troponin I (slow) promoters by about 2–3-fold in the presence of calcineurin. Two of the known effectors of calcineurin signalling on gene expression in muscle are the Nfat and Mef2 transcription factor families^{20–23}. To determine whether PGC-1 α exerts its effects on type I promoters by co-activating these transcription factors, the troponin I (slow) promoter was co-transfected with Nfat or Mef2, with or without PGC-1 α . As shown in Fig. 3b, PGC-1 α had little effect on the transcriptional activity of Nfatc3, and only minimal effect on Mef2a, whereas it co-activated Mef2c and Mef2d by about 3- and 6-fold, respectively²⁴. These results are consistent with our previous report that PGC-1 α physically interacts with Mef2c and co-activates this protein on the Glut-4 promoter in muscle cells²⁵.

To more critically assess the function of the Mef2 factors in the role of regulation of oxidative muscle gene expression by PGC-1 α , we used a well-characterized version of the myoglobin promoter (–373 to +7) containing a mutation in the Mef2 binding site⁹. PGC-1 α is capable of activating the wild-type promoter in the presence of calcineurin or Mef2d. The transcriptional activity was slightly higher than the combination of calcineurin and PGC-1 α when all three factors were added (data not shown). As expected, the ability of PGC-1 α to co-activate this promoter in combination with Mef2d was greatly diminished in the promoter lacking a functional

Mef2 site. Significantly, transcription from this slow-type promoter was reduced in the absence of the Mef2 binding site in response to calcineurin and Mef2d (Fig. 3c). As suggested by these functional data, PGC-1 α interacts directly with several Mef2 proteins²⁵ but does not seem to co-activate the transcriptional activity of Nfatc3. PGC-1 α increased the activity of the Mef2 mutant promoter by about twofold in the presence of calcineurin (Fig. 3c), suggesting that other transcription factors downstream of calcineurin may also be important for PGC-1 α co-activation. Taken together, these results suggest that cooperation between PGC-1 α and the calcineurin pathway is probably critically important in switching of muscle fibre types or myofibre transformation to oxidative muscle, and a significant portion of this interaction may occur as a consequence of the direct co-activation of Mef2 proteins by PGC-1 α .

Previous studies of muscle fibre specification indicated an important role for exercise, electrical activity, calcium fluxes and calcineurin in causing conversion of type II to type I fibres. At a transcriptional level, analysis of slow compared with fast muscle fibres had indicated an important role for Nfat and Mef2 proteins in the activation of slow-fibre-selective myofibrillar proteins^{21,26}. However, there had not been described to date a transcriptional component that could potentially integrate calcium signalling, mitochondrial biogenesis and these myofibrillar protein regulators. The studies presented here indicate that PGC-1 α is a principal physiological regulator (a switch) for type I fibre specification. There are several key pieces of data supporting this. First, PGC-1 α is expressed preferentially in muscles rich in type I fibres, such as the soleus. Second, elevation of PGC-1 α expression in muscles rich in type II fibres evokes marked changes in muscle morphology, gene expression and function. Whereas the role of PGC-1 α in activating the genes of mitochondrial oxidative metabolism in muscles *in vivo* was not surprising, a role for PGC-1 α in stimulating the expression of myofibrillar proteins characteristic of type I muscle is entirely unanticipated. The data showing the reduced fatigability of muscle

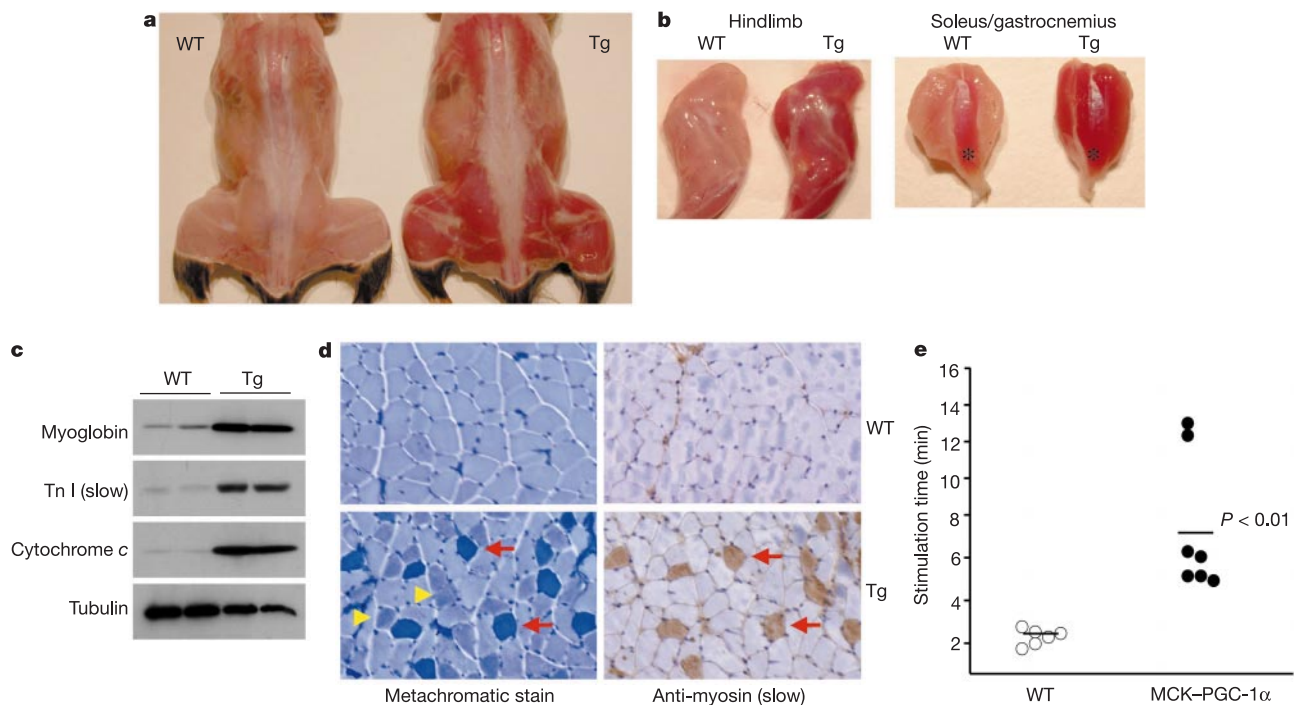


Figure 2 Morphological, histological and functional properties of skeletal muscles of the MCK-PGC-1 α transgenic mouse (strain 31). **a**, Dorsal view of wild-type (WT) and transgenic (Tg) animals. **b**, Morphology of dissected hindlimb and gastrocnemius/soleus (asterisk) muscle. **c**, Immunoblotting analyses of various proteins in white vastus muscle

of transgenic mice and wild-type controls. **d**, Metachromatic¹⁹ and immunohistochemical staining of wild-type and transgenic plantaris muscle. Arrows indicate type I fibres whereas arrowheads point to type IIa fibres. **e**, Fatigue index of EDL muscle (see text). Data were analysed by one-way analysis of variance (ANOVA) with *P*-value indicated.

from the transgenic mice indicates that this represents a genuine and functional conversion of myofibres towards a more oxidative state. That this conversion probably represents an authentic biological role of PGC-1 α is strongly suggested by the fact that these changes are all brought about through levels of PGC-1 α protein equal to those seen in type I muscle (Fig. 1c). The fibre type conversion driven by PGC-1 α does not completely convert all fibres, suggesting that other signals also influence fibre phenotype. In this regard, fibre type conversion driven by PGC-1 α is similar to what has been observed with calcineurin transgenics and exercise^{10,21}.

Previous data had shown that PGC-1 α mRNA levels can be increased by exercise in rats²⁷, but how physical activity and/or nerve signals might control PGC-1 α is unknown. Given the estab-

lished role for calcium signalling in fibre type conversion and the preferential expression of PGC-1 α in muscle rich in type I fibres, it is likely that PGC-1 α is downstream of the calcium/calcineurin/CaM kinase pathway (Fig. 3d). Alternatively, muscle contraction results in increased energy demand and activation of the cellular energy sensing pathway, for example through increasing the activity of AMP-activated kinase⁶, which could also lead to the induction of PGC-1 α expression. As shown previously, PGC-1 α powerfully co-activates nuclear respiratory factors and activates the programme of mitochondrial biogenesis and fuel oxidation. As shown here, PGC-1 α also increases the expression of type I fibrillar proteins by co-activating Mef2 transcription factors, thereby coordinating the expression of both metabolic and contractile properties of type I myofibres. Although Mef2 and Nfat have been shown to activate gene expression of type I fibres⁹, no functional interaction between PGC-1 α and Nfat on troponin I (slow) promoter was observed in the current study.

These results might be useful ultimately in allowing improved control of muscle fibre types, physiology and metabolism. Although we have not yet examined the effects of transgenic PGC-1 α on chronic glucose and energy metabolism, the rate of glucose uptake and oxidation in muscle is thought to be a critical rate-limiting step in whole-body glucose disposal under the influence of insulin. Hence, pharmacological manipulation of PGC-1 α and the principal docking events in fibre-type specification could have implications for obesity and diabetes. □

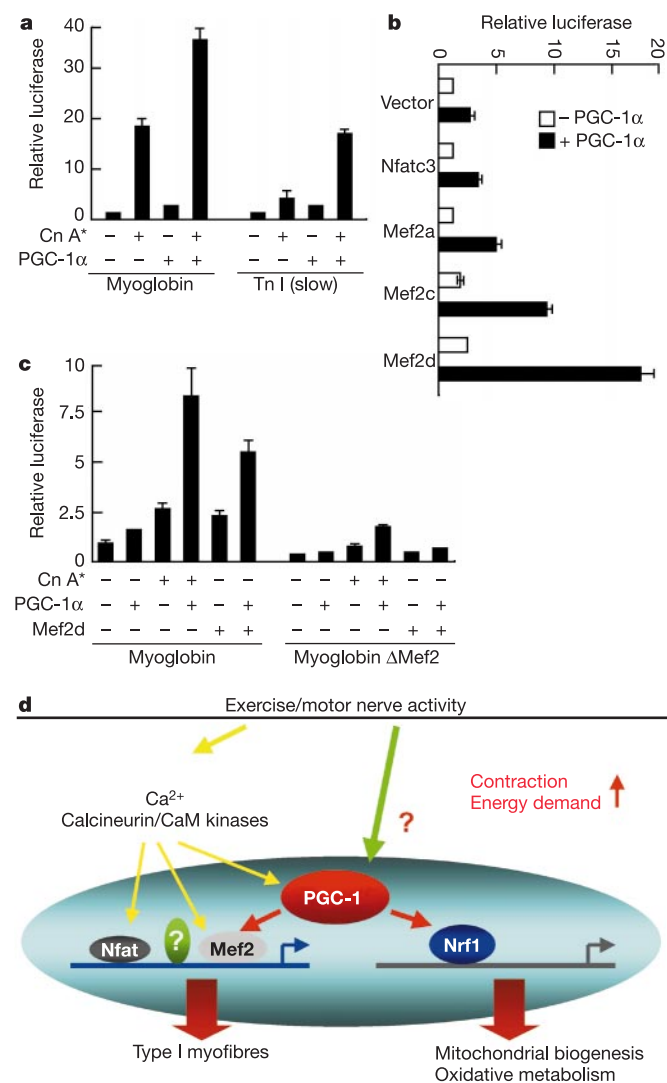


Figure 3 Co-activation of the calcineurin signalling pathway by PGC-1 α . **a**, PGC-1 α co-activates the calcineurin signalling pathway. C2C12 myoblasts were transiently transfected with myoglobin or troponin I (slow) promoter luciferase constructs along with constitutively active calcineurin (Cn A*)⁹, PGC-1 α , or the combination as indicated. Luciferase activity was normalized to the basal activity of respective promoters. **b**, Functional analysis of Nfat and Mef2 proteins with PGC-1 α on the troponin I (slow) promoter. **c**, Requirement for a proximal Mef2 binding site in PGC-1 α co-activation of the myoglobin promoter. Wild-type or Mef2 mutant (Δ Mef2) myoglobin promoter luciferase constructs⁹ were transiently transfected into C2C12 myoblasts as indicated. **d**, A model for the role of PGC-1 α in muscle fibre specification.

Methods

Generation of transgenic mice and analyses of gene expression

PGC-1 α complementary DNA was placed downstream of a 6.5-kilobase (kb) MCK promoter sequence¹⁷. MCK-PGC-1 α transgenic mice were generated by standard DNA microinjection and identified by PCR-based genotyping. For gene expression analyses, muscles were dissected from 3-month-old wild-type or transgenic C57/Bl6 mice. All experiments were performed in transgenic line 31 except where indicated. RNAs were isolated using Trizol (Gibco BRL) and analysed by hybridization with specific cDNA probes. Protein lysates were prepared in a buffer containing PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and a cocktail of protease inhibitors, separated by SDS-polyacrylamide gel electrophoresis, and immunoblotted with specific antibodies.

Metachromatic and immunohistochemical staining

Skeletal muscle was collected, embedded in gum tragacanth mixed with OCT freezing matrix, and quickly frozen in isopentane cooled in liquid nitrogen. Eight-micrometre-thick serial sections were obtained and processed for immunohistochemical staining with a monoclonal antibody against skeletal slow myosin (NOQ7.5.4D, Sigma). Fibre typing was performed with cryostat sections of plantaris muscles using metachromatic dye-ATPase methods as described¹⁸.

Measurements of muscle fatigue

Wild-type or transgenic mice were killed and EDL muscles were immediately dissected and mounted between a fixed clamp at the base of a water-jacketed organ bath and a Grass FTO3C isometric force transducer. Muscles were equilibrated in an oxygenated (95% O₂ and 5% CO₂) physiological salt solution at 30 °C for 15 min and the resting force was maintained at 1 gram¹⁹. Muscles were then stimulated at 100 Hz until the force output drops to 30% of the initial level. The time taken to reach 30% of the original force was measured and used as an indication for muscle fatigability.

Cell culture and transient transfection

C2C12 myoblasts were maintained in DMEM in the presence of 10% FBS. Transient transfection was performed at a confluence of approximately 75% using Superfect (Qiagen) with 0.5 μ g reporter plasmid with or without other factors (1.0 μ g). Lysates from transfected cells were prepared 24–36 h after transfection and assayed for luciferase activity. All transfection experiments were repeated at least three times. Myoglobin and troponin I (slow) promoter luciferase plasmids were previously described⁹.

Received 19 April; accepted 29 May 2002; doi:10.1038/nature00904.

- Berchtold, M. W., Brinkmeier, H. & Muntener, M. Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiol. Rev.* **80**, 1215–1265 (2000).
- Puigserver, P. et al. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
- Wu, Z. et al. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
- Vega, R. B., Huss, J. M. & Kelly, D. P. The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor α in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Mol. Cell. Biol.* **20**, 1868–1876 (2000).

5. Booth, F. W. & Thomason, D. B. Molecular and cellular adaptation of muscle in response to exercise: perspectives of various models. *Physiol. Rev.* **71**, 541–585 (1991).
6. Hood, D. A. Contractile activity-induced mitochondrial biogenesis in skeletal muscle. *J. Appl. Physiol.* **90**, 1137–1157 (2001).
7. Gollnick, P. D. *et al.* Enzyme activity and fibre composition in skeletal muscle of untrained and trained men. *J. Appl. Physiol.* **33**, 312–319 (1972).
8. Jarvis, J. C. *et al.* Fast-to-slow transformation in stimulated rat muscle. *Muscle Nerve* **19**, 1469–1475 (1996).
9. Chin, E. R. *et al.* A calcineurin-dependent transcriptional pathway controls skeletal muscle fibre type. *Genes Dev.* **12**, 2499–2509 (1998).
10. Naya, F. J. *et al.* Stimulation of slow skeletal muscle fibre gene expression by calcineurin *in vivo*. *J. Biol. Chem.* **275**, 4545–4548 (2000).
11. Bigard, X. *et al.* Calcineurin co-regulates contractile and metabolic components of slow muscle phenotype. *J. Biol. Chem.* **275**, 19653–19660 (2000).
12. Olson, E. N. & Williams, R. S. Remodeling muscles with calcineurin. *Bioassays* **22**, 510–519 (2000).
13. Lehman, J. J. *et al.* Peroxisome proliferator-activated receptor γ coactivator-1 promotes cardiac mitochondrial biogenesis. *J. Clin. Invest.* **106**, 847–856 (2000).
14. Yoon, J. C. *et al.* Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001).
15. Esterbauer, H. *et al.* Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression. *Genomics* **62**, 98–102 (1999).
16. Lin, J. *et al.* PGC-1 β : a novel PGC-1 related transcription coactivator associated with host cell factor. *J. Biol. Chem.* **277**, 1645–1648 (2002).
17. Johnson, J. E., Wold, B. J. & Hauschka, S. D. Muscle creatine kinase sequence elements regulating skeletal and cardiac muscle expression in transgenic mice. *Mol. Cell. Biol.* **9**, 3393–3399 (1989).
18. Ogilvie, R. W. & Feedback, D. L. A metachromatic dye-ATPase method for the simultaneous identification of skeletal muscle fibre types I, IIA, IIB and IIC. *Stain Technol.* **65**, 231–241 (1990).
19. Grange, R. W. *et al.* Functional and molecular adaptations in skeletal muscle of myoglobin-mutant mice. *Am. J. Physiol. Cell Physiol.* **281**, C1487–C1494 (2001).
20. Wu, H. *et al.* MEF2 responds to multiple calcium-regulated signals in the control of skeletal muscle fibre type. *EMBO J.* **19**, 1963–1973 (2000).
21. Wu, H. *et al.* Activation of MEF2 by muscle activity is mediated through a calcineurin-dependent pathway. *EMBO J.* **20**, 6414–6423 (2001).
22. Calvo, S. *et al.* Fibre-type-specific transcription of the troponin I slow gene is regulated by multiple elements. *Mol. Cell. Biol.* **19**, 515–525 (1999).
23. Black, B. L. & Olson, E. N. Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu. Rev. Cell Dev. Biol.* **14**, 167–196 (1998).
24. Delling, U. *et al.* A calcineurin-NFATc3-dependent pathway regulates skeletal muscle differentiation and slow myosin heavy-chain expression. *Mol. Cell. Biol.* **20**, 6600–6611 (2000).
25. Michael, L. F. *et al.* Restoration of insulin-sensitive glucose transporter (Glut4) gene expression in muscle cells by the transcriptional coactivator PGC-1. *Proc. Natl Acad. Sci. USA* **98**, 3820–3825 (2001).
26. Nakayama, M. *et al.* Common core sequences are found in skeletal muscle slow- and fast-fibre-type-specific regulatory elements. *Mol. Cell. Biol.* **16**, 2408–2417 (1996).
27. Goto, M. *et al.* cDNA cloning and mRNA analysis of PGC-1 in epitrochlear muscle in swimming-exercised rats. *Biochem. Biophys. Res. Comm.* **274**, 350–354 (2000).

Acknowledgements

We thank R. Sanders Williams and A. Buonanno for providing reporter constructs, and J. Lawitts for generating transgenic mice. This work was supported by grants from the NIH

to B.M.S., B.B.L. and E.N.O. J.L. was supported by a postdoctoral fellowship from the American Heart Association.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.M.S. (e-mail: bruce_spiegelman@dfci.harvard.edu).

erratum

A new hominid from the Upper Miocene of Chad, Central Africa

Michel Brunet, Franck Guy, David Pilbeam, Hassane Taisso Mackaye, Andossa Likius, Djimdoumalbaye Ahounta, Alain Beauvilain, Cécile Blondel, Hervé Bocherens, Jean-Renaud Boisserie, Louis De Bonis, Yves Coppens, Jean Dejax, Christiane Denys, Philippe Douring, Véra Eisenmann, Gongdibé Fanone, Pierre Fronty, Denis Geraads, Thomas Lehmann, Fabrice Lihoreau, Antoine Louchart, Adoum Mahamat, Gildas Merceron, Guy Mouchelin, Olga Otero, Pablo Pelaez Campomanes, Marcia Ponce De Leon, Jean-Claude Rage, Michel Sapanet, Mathieu Schuster, Jean Sudre, Pascal Tassy, Xavier Valentin, Patrick Vignaud, Laurent Viriot, Antoine Zazze & Christoph Zeltkefer

Nature **419**, 145–151 (2002).

Owing to an editorial oversight, the sentence in the second column of page 150 that reads ‘Lower c and the lower and upper premolars each have three pulp canals and two roots’ is in error. Instead, it should read ‘Lower and upper premolars each have three pulp canals and two roots’.

Degrees of separation

The first of two tranches of new chromatography products.

Horizon Pioneer

Biotage

www.biotage.com

HPFC starter kit

Biotage's high-performance flash chromatography technology (HPFC) powers the company's Horizon purification system. The Horizon Pioneer has been designed to provide basic HPFC. It can be upgraded to a full Horizon system by adding fraction collection, ultraviolet detection, or both. The Pioneer has a pump capable of generating pressures exceeding 100 psi and can run at flow rates of up to 100 ml per minute. A multisegmented gradient eliminates the need to pre-mix solvents, and a digital display shows real-time gradient composition and flow rate.

200 Series

Cambridge Scientific Instruments

www.camsci.co.uk

Electronic pressure control

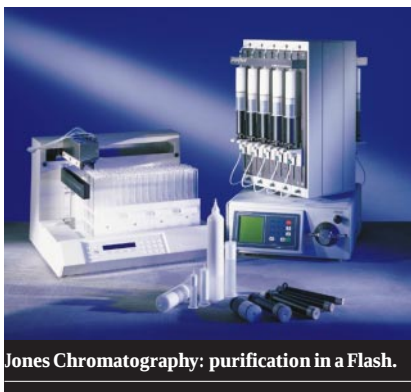
The 200 series gas chromatograph from Cambridge Scientific features electronic pressure control (EPS) as standard. EPS eliminates needle valves and mass flow controllers, improves resolution, reduces analysis time and extends column life, since lower column temperatures can be used. Other features include manual or automatic gas sampling and an internal 'timed event' facility that can be used to control unattended operation or to change parameters during a run.

FlashMaster Parallel

Jones Chromatography www.jones-chrom.co.uk

Binary gradient capability

The FlashMaster Parallel chromatography system is designed to speed up throughput in parallel organic synthesis purification. It offers parallel purification of up to 10 samples, a disposable column system and fraction collection. The binary gradient function allows



Jones Chromatography: purification in a Flash.

the creation of multi-segment step and linear gradients to cope with difficult separations. The system is especially useful for coupling to commercial parallel organic synthesizers for on-line synthesis purification.

Chromatography columns

Omnifit

www.omnifit.com

Bores come in all sizes

These columns are available with bores of 3–50 mm, and allow precision liquid chromatography to be carried out in the laboratory. Made of borosilicate glass with PTFE fittings, the columns have minimal dead volume to eliminate sample contamination and ensure smooth solvent flow. All columns and attachments are designed for simple assembly and disassembly. They are suitable for a variety of analytical, preparative and high-performance applications as well as original equipment manufacture.

821 Compact

Metrohm

www.metrohm.co.uk

Affordable on-line chromatography

Metrohm's 821 compact on-line ion chromatography system is a simplified version of its more expensive 811 system. The 821 can make sequential determination of anions or cations from up to 10 sample flows. The system contains a PC and a modem for remote maintenance, and software with control cards for measurement results, calibrations and performance checks. It comes with a removable trolley or can be wall-mounted. The 821 needs only monthly maintenance and can be operated by non-skilled chemists with basic training.

Chromegabond HPLC columns

ES Industries

www.esind.com

Quality, on the level

ES Industries describes its Chromegabond columns for high-performance liquid chromatography (HPLC) as 'quality defined', promising optimization of spherical silica, bonding reagents and procedures, batch testing, column packing and final column testing. The columns can be used at all levels of analysis, including method development, liquid chromatography/mass spectrometry analysis and quality control.

Cyclograph II

Analtech

www.analtech.com

Get out of a scrape

The Cyclograph II chromatography system from Analtech uses centrifugal force to pro-



Omnifit columns, pure and simple.

duce thin-layer radial separations. The compound to be separated is applied inside the adsorbent ring of the pre-cast rotor and the selected solvent mixture is then pumped through the adsorbent layer. The centrifugal force of the rotor accelerates the normally slow capillary action of the solvent through the layer and fractions are collected within about 20 minutes. There is no need to scrape separated bands from the layer. The system features a built-in ultraviolet lamp, hinged lid, touch pad control system and drop-in rotor mounting.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Silke Opstrup (4994)

France/ Belgium:

Amelie Pequignot (4974)

Production Manager:

Billie Franklin
To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Switzerland:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Harakakatamachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746
e-mail: k.johnson@naturejp.com

Asia-Pacific Advertising Manager:

Keyn Johnson

naturejobs

A decisive effort?

The UK government last month announced plans to boost the Department of Trade and Industry's budget from £2 billion (US\$3 billion) to £2.9 billion in 2005–6. The proposals for the department, which supports the bulk of UK university research, certainly caught the attention of the scientific establishment (see *Nature* **418**, 261; 2002). But the question remains, will it capture the imagination of young students who are undecided about whether to pursue science or a different career with a faster, bigger financial payoff?

In the past few years, these 'fence-sitters' seemed to have jumped, with fewer pursuing scientific careers — especially in physics and chemistry (see *Nature* **416**, 777; 2002). But given years of low stipends for both graduate students and postdocs, coupled with the recent new burden of student loans, it might be more accurate to say that they have been pushed. The new proposal might help to push them in the other direction, into the public science arena.

The funding promise provides both long- and short-term incentives. A bigger overall science budget means more money for scientific overheads, which would make prospective graduate students feel that they have a future once they secure their PhD. New funding to help researchers spin their work off into companies also helps. But the short-term incentives might be more noticeable to potential science graduate students.

Under the plan, PhD grants would grow from £10,000 to £13,000 a year, and postdoc stipends would rise by £4,000. As a percentage these numbers are massive, as a reality less so. Perhaps they could be made more enticing by coming into effect immediately, rather than incrementally. A fast hike could attract prospective young scientists' attention, whereas a slow one, which over time might not be much more than a modest cost-of-living boost, would just prompt them to jump the other way.

Paul Smaglik
Naturejobs editor



Contents

CAREERS AND RECRUITMENT

Nanotechnology finds its feet p4

Japan builds bridges between biology and physics p7

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

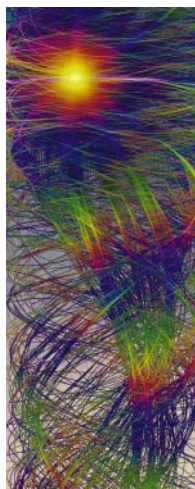
FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

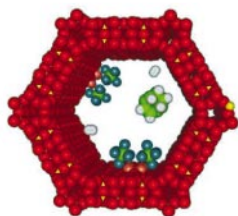
SCIENTIFIC EVENTS



Small world, big opportunities

At last nanotechnology is moving from the realm of hype and hope into the real world, with jobs and funding appearing on both sides of the Atlantic. Paul Smaglik considers the options.

Nanotech inaction: mixed-metal clusters anchored within MCM-41 silica for use in catalysis.



Tiny detective: this microsensor, shown in front of a fly's compound eye, can detect biological and chemical molecules with extreme sensitivity.



Nanotechnology is beginning to bridge the gap between hype and reality. The word may have been in use for years, but it has never really been properly defined. Few institutes actually specialized in the area, and the kind of massive funding increases that can revolutionize a field and create many new jobs were only promises.

Now that the money is materializing, nanotechnology is being seen as simply another name for molecular-scale physics, chemistry and biology — although it is true that the convergence of these three disciplines has also created something new. Yet the size of the field has increased dramatically. In the United States, the National Nanotechnology Initiative (NNI) will bring it more than \$700 million next year. The US National Science Foundation has created six institutes, the Department of Energy has unveiled

plans for five more, and the Department of Defense and NASA are also funding significant projects.

Europe, too, is taking action. The sixth Framework funding programme being planned by the European Commission (EC) stands to pledge up to 1.2 billion euros (US\$1.2 billion) over the next four years to establish nanotech research networks. This is in addition to member countries' own programmes, which are also growing, especially in France, the United Kingdom and Germany.

Researchers wanting to obtain some of that money — and make the best use of it — could take advice from Mike Roco, chairman of the NNI, and Ramón Compañó, scientific officer at the European Commission in charge of nanotech funding. They recommend finding a lab that specializes in nanotechnology, especially if it has been elevated to the status of a 'centre'. Then, they say, learn as many skills from disciplines other than your own as possible and foster research relationships with other labs.

Picks and shovels

Businesses that might create the most nanotech jobs initially will be the ones providing the 'picks and shovels' for the field, says Ramón Compañó, European Commission (EC) scientific officer for nanotechnology funding. These include manufacturers of

instruments such as scanning tunnelling microscopes (see *Nature* 410, 127–130; 2001) and providers of materials such as nanoparticles.

But picking a winning technology — much less a specific company — is not so easy, he says.

For example, the EC recently funded two projects on nano-imprint lithography, in the hope that this would boost the semiconductor industry. But instead, he says, the technology has been co-opted for biological or optical applications by companies such as Obducat of Malmö in Sweden, Mildendo of Jena in Germany and the EV Group of Austria.

P.S.

YOUTH APPEAL

"A major bottleneck to making the nano-revolution happen is people," Compañó says. "We need many young talented people, and these people have to be trained. It takes quite a long time to give them a solid formation in one of the traditional disciplines and to get them familiar with others." This is doubly challenging because fewer people are being drawn into chemistry and physics than in the past.

Patricia Dehmer, an associate director of basic research at the US Department of Energy, agrees that training is a challenge. "You either have to know something about multiple disciplines yourself or be really good at putting together coalitions," Dehmer says. But she thinks that the raised profile of nanotechnology — and its increasing budget — will change the pipeline problem. "Students go where the excitement is and where the funding is."

Early reports from the NNI suggest that this is already happening in the United States. Chad Mirkin, who runs a nanotech centre at Northwestern University in Evanston, Illinois, says that being designated as one of the National Science

A glimpse of the future

Scientists wondering what research organizations will look like under the European Commission's sixth Framework programme, which is set to begin next year, would do well to check out Phantomsnet. This is a programme that provides a virtual coalition of research groups studying nanoelectronics, and has been around for almost two years.

Membership has grown from 24 groups to 180, with about 10 new applications every few weeks. Research groups must gain their own funding, but Phantomsnet helps them to interact with each other by

sponsoring conferences, hosting a website which now receives 2.5 million hits a month, and underwriting scientific exchange, including sending some members to US conferences and bringing South American researchers to European labs.

Its next project is to increase industrial membership, which has grown from a handful of members to about 13% of the total membership. But perhaps the most important goal is to stay in existence: Phantomsnet will need to reapply for funding under the sixth Framework programme if it is to stay in operation for the next four years. P.S. www.phantomsnet.com

Small thoughts: a crossed junction of nanotubes formed by self-assembly on a silanized template.

Foundation's six nanotech hubs last year is already paying off. "This year we recruited 72 graduate students," he says. "Our average is 35."

Mirkin expects recruiting to be even easier now following the opening this month of the centre's new facility dedicated to patterning biological molecules. The centre's emphasis on the intersection of biology and physics should help it to attract young scientists. "Our main focus and our main strength is the medical aspect of nanotechnology," he says. He hopes that the centre can help to develop medical diagnostics that can exploit recent advances in genomics and proteomics.

Mirkin says that nanotech's increasing reputation has been useful to his students. "I've never had a student have a problem getting a job," he says. He receives regular calls from recruiters looking for people with experience in synthesis. His most common response is: "You're way too late."

TRAINING WAYS

Angela Belcher, who this autumn will move her University of Texas lab to the Massachusetts Institute of Technology, unintentionally followed an ideal educational path to her nanotechnology career. Her first degree emphasized biochemistry, whereas her PhD was in inorganic chemistry. Then she rounded things out with a postdoc in electrical engineering.

She looks for graduate students and postdocs who have similarly diverse backgrounds, or who aren't averse to developing them. "I select for people who are really interested in expanding their knowledge base," Belcher says. "People who are more afraid to do new kinds of science or new kinds of disciplines wouldn't fit in well."

Her goal is to teach everyone in her lab the same basic skill sets, so

that it would be hard to tell what discipline they came from. "Everyone in my lab can clone and sequence DNA, do inorganic synthesis and characterize materials using various imaging techniques."

That multidisciplinary approach has paid off. Belcher's lab has had some success in inducing viruses and bacteria to produce inorganic materials. She next plans to work in the other direction, using inorganic materials to probe cellular functions. A growing number of companies are interested in exploring the connection between biology and material sciences, she comments.

Although Belcher's career illustrates a direct path to nanotechnology, Carlo Montemagno's shows that detours can still take you to that destination — in his case, to the California NanoSystems Institute.

With a first degree in biological engineering and a PhD in civil engineering, Montemagno spent most of his early career in microscale-fluid physics. But reading multidisciplinary literature in the mid-1990s gave him some new ideas. "I was able to identify that there was a convergence of technology occurring in the physical and life sciences," Montemagno says. So while at Cornell University, he started teaching himself about molecular biology.

Apart from reading the literature, he took nine months off from his research to work side-by-side with life-science postdocs in colleagues' labs and to visit other universities. From there, he had enough of a skill set to begin research on molecular motors. But he first needed to change laboratories and get new equipment.

He admits the changeover was a risk, but says that the process of re-educating himself mid-career was "very exhilarating". And the risk has paid off — it led in 2001 to a chair in biomedical engineering at the

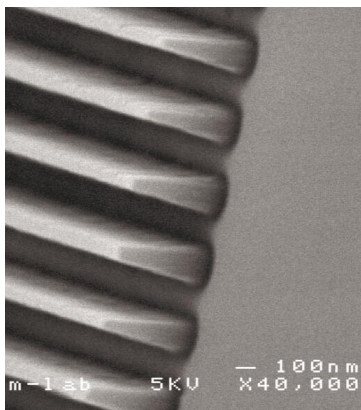
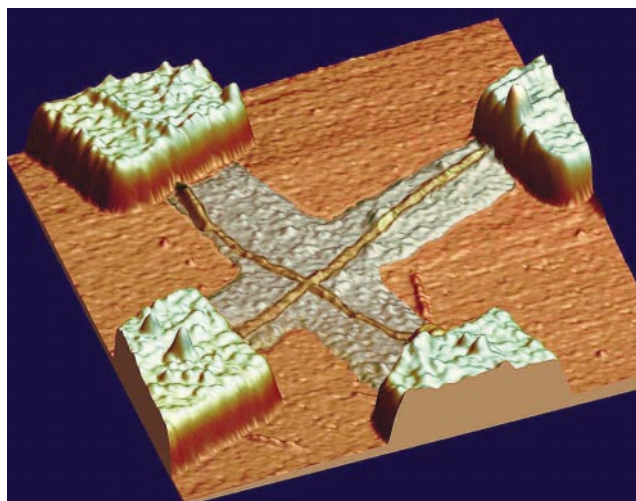


Angela Belcher (above) is taking a multidisciplinary approach to nanotechnology.

Making a mark: nanoimprint lithography stamps (opposite).

CEA-MOTOROLA LEM

C. CARDINAUD/IMN/CNRS



University of California, Los Angeles, where he could develop a programme in bioengineering from scratch.

Although affiliation with one of the large centralized labs is an advantage in the United States, nanotech success in Europe is more likely to grow from membership of a diffuse network.

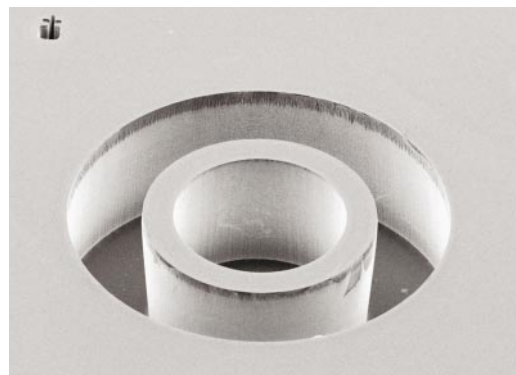
The promised funding from the EC's sixth Framework programme has led some 15,000 groups to express an interest in taking part — from all disciplines, some not actually in nanotech — although only a few hundred are likely to share in the funding. But even for those whose applications are unsuccessful, the exercise may still pay off, spurring collaborations that could be fruitful even though they are unfunded.

Compañó sees collaborations creating virtual centres that span Europe. For example, he envisages a virtual nanoelectronics centre being formed from labs throughout Europe using different areas of expertise — engineers who understand circuits, chemists who can synthesize molecules, and biologists who could perhaps fathom how DNA could act as wiring. An informal network, Phantomsnet, already exists (see 'A glimpse of the future', page 5).

NANO NETWORKS

Christian Joachim and Gunter Reiter, directors of two French nanotech labs, are following different strategies in competing for this pot of EC gold. Joachim, of the Centre for the Study and Structural Elaboration of Materials in Toulouse, hopes to establish a formal network including five or six groups he is in contact with and whose research relationships are already defined.

Reiter, of the Institute of Chemistry at Surfaces and Interfaces in Mulhouse, is casting a wider net, hoping to attract 50–60 collaborators from across Europe. This October he plans to host a meeting in Crete



where researchers in all relevant disciplines can find out more about the work that is being done, and clarify their own role, before making formal applications for funding.

To compete for the money, many labs that would normally compete with each other for smaller grants are banding together in the hope of sharing a larger grant. "Everybody talks about who is in which project and which project has the best chances," Reiter says. He feels excited about the process, because it could make European nanotech a more international endeavour. An Austrian, Reiter has worked in Germany as well as France. He hopes to foster relationships between neighbouring labs in France and Germany, and with others that are not so physically close.

"One of the things I realized when I started to work in France was that it was really difficult to get people to realize that it is possible to collaborate with Germany," Reiter says. "Things are changing now." He is hearing more bilateral calls for proposals from the French and German governments and has heard hints that the DFG, Germany's main research funding agency, might start supporting more collaborative arrangements between Germany and other countries.

The downside, professionally, is time. Reiter anticipates that the EC application process will be long and cumbersome. He says potential collaborators are asking: "Is it really necessary to do this extra work to write these highly competitive proposals?"

Joachim asks the same question, but is more pessimistic about the overall odds. "If you think that when you propose you will be accepted, you will be very disappointed," he says.

He is uncertain whether the programme will result in a net gain for nanotechnology, as he is not sure if the new framework will add to or simply replace efforts funded in the previous one. He has already seen plenty of shuffling of resources within France. For example, although his lab has grown from three permanent scientists 10 years ago to 16 now, those positions have come from losses in other labs.

Still, like many of his colleagues, Joachim will heed the EC's call and look for more collaborators outside France to boost his odds. "Without other countries, you're dead," he says.

Paul Smaglik is editor of Naturejobs.

US National Nanotechnology Initiative ♦ www.nano.gov

The sixth Framework programme and nanotechnology

♦ www.cordis.lu/rtd2002/fp-activities/nanotechnologies.htm

National Academy of Science's nanotechnology report

♦ www.nap.edu/catalog/10395.html

Nanotech shuffle

Nanotech collaborations are springing up in Britain, including the £9-million (US\$13.8-million) project linking the University of Cambridge, University College London (UCL) and the University of Bristol. But will they result in more jobs? UCL physics professor Marshall Stoneham is unsure.

Many such arrangements are reorganizations of existing resources, he says: "There's been a lot of shuffling around." The tri-university collaboration will at least get its own new building.

On the other hand, Stoneham notes, the movement may benefit young scientists. "At the moment there are quite a lot of postdoc opportunities," he says.

But he cautions that taking such positions may amount to a calculated risk — until real positions are created in academia or industry (see 'Picks and shovels' page 4). The problem is the long term, he says, as "the career structure is not actually established". **P.S.**

Interdisciplinary Research Collaboration in Nanotechnology

♦ www.nanoscience.cam.ac.uk/irc

Into the miniature world: a 1-mm sensor made from silicon to detect the high-speed changes in pressure inside a turbine (top right); and the surface structure of iron silicide showing ordering of the atomic layers (above).

Breaking down biological borders

Opportunities in nanotechnology are opening up in Japan — especially for young researchers willing to cooperate across disciplines, says Robert Triendl.

Japan's love affair with nanotechnology has, until recently, tended to centre on physics. Most of Japan's nanotech funding, which has been supported by a number of agencies since the 1980s, has remained oriented towards studying nanoscale phenomena in semiconductor materials or developing new materials.

But that is changing. In recent months, policy-makers have been strengthening research capacity at the interface between life sciences and nanotechnology. Research centres in the public sector are now increasingly offering positions for scientists specializing in nanobiology. And, perhaps most importantly, funding schemes for young scientists in nanobiotechnology are emerging.

YOUTH MOVEMENT

The Japan Science and Technology Corporation funds emerging research areas through schemes such as ERATO, CREST and PRESTO. ERATO (Exploratory Research for Advanced Technology) provides relatively young scientists with a sizeable amount of money — typically between US\$15 million and US\$18 million over five years — and almost total freedom to spend it.

Toshio Yanagida and his group at Osaka University are one of ERATO's success stories. The experimental techniques they designed during their studies of how 'biomotors' work inside muscle cells have helped to revolutionize the fields of single-molecule imaging and the study of motor proteins. Today, some five years

after the project ended, many of the young scientists hired by Yanagida have moved to universities or research institutes across Japan, taking with them knowledge of the techniques developed in Yanagida's lab.

The development of such techniques appears to be one of the brightest areas for employment opportunities in nanobiotechnology. At

the Nanotechnology Research Institute (NRI) in Tsukuba, the main research focus is on molecular nanotechnology for applications such as molecular devices or drug-delivery systems. Hiroshi Yokoyama, the NRI's director, says that he plans to hire more biologists. Other institutes, including the new Tissue Engineering Research Center in Osaka, are also strengthening interdisciplinary research in nanobiology.

According to Yokoyama, some 10% of research funding at his institute is spent on the development of new research tools. "Instrumentation tools provide the basic infrastructure for research in nanotechnology — and a platform for cooperation," he says.

Satoshi Kawata, a professor of applied physics at Osaka University, agrees that instrumentation is crucial for nanotechnology. Many tools used by nanotech scientists, such as the atomic force microscope, were originally developed by physicists to study fundamental physical phenomena or the physical properties of new materials. But, says Kawata, it is equally important to develop organizational models that allow cooperation between scientists designing the new techniques and those who will use them, such as cell biologists.

ORGANIZING COOPERATION

At universities, venture-business laboratories and interdisciplinary research centres are offering space for the development of cooperative endeavours by researchers from various disciplines — although sometimes with mixed results. Osaka University's Frontier Research Center goes a step further.

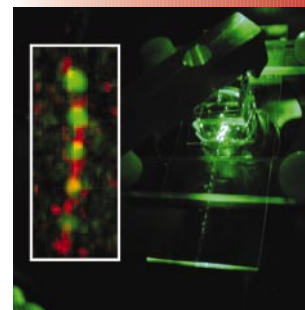
Set up late last year with government support, the centre aims to provide a better organizational mechanism to fund high-quality research in rapidly emerging areas and to aid cooperation with industry.

But despite such efforts, cooperation between physicists and biologists remains limited. According to Yokoyama, the barriers between disciplines in the nanotechnology field are only slowly breaking down. "Japan's academic circles remain very closed," he says. ■

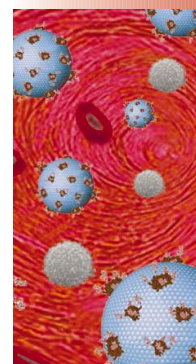
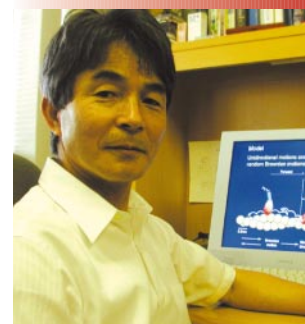
Robert Triendl is a freelance writer based in Tokyo.

ERATO ♦ www.jst.go.jp/erato

Japan Science and Technology Corporation ♦ www.jst.go.jp/EN
Nanotechnology Research Institute ♦ unit.aist.go.jp/nanotech
Tissue Engineering Research Institute ♦ unit.aist.go.jp/terc
Frontier Research Center ♦ www.frc.handai.com



Toshio Yanagida (below) is a pioneer in techniques and instruments (above) for imaging single molecules such as myosin (inset).



On target: this drug-delivery system, designed in Hiroshi Yokoyama's lab, uses sugar-protein conjugates as receptors.



Hiroshi Yokoyama (opposite) sees instrumentation as a key area of growth for nanotechnology.